UNIVERSIDADE FEDERAL DE MINAS GERAIS

Faculdade de Ciências Econômicas (FACE)

Centro de Desenvolvimento e Planejamento Regional (Cedeplar)

Programa de Pós-graduação em Economia

Victor Medeiros

Empirical Essays on Infrastructure and Regional Development:

Bringing highway investments more efficient, inclusive, and sustainable

Belo Horizonte

2024

Victor Medeiros

EMPIRICAL ESSAYS ON INFRASTRUCTURE AND REGIONAL DEVELOPMENT:

Bringing highway investments more efficient, inclusive, and sustainable

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Economia.

Orientador: Prof. Dr. Rafael Saulo Marques Ribeiro.

Coorientador: Prof. Dr. Pedro Vasconcelos Maia do Amaral.

Belo Horizonte

2024

Ficha catalográfica

M488e 2024

Medeiros, Victor.

Empirical essays on infrastructure and regional development [manuscrito]: bringing highway investments more efficient, inclusive, and sustainable / Victor Medeiros. – 2024.

1v.: il.

Orientador: Rafael Saulo Marques Ribeiro.

Coorientador: Pedro Vasconcelos Maia do Amaral.

Tese (doutorado) – Universidade Federal de Minas Gerais, Centro de Desenvolvimento e Planejamento Regional.

Inclui bibliografia.

1. Infraestrutura (Economia) - Teses. 2. Desenvolvimento econômico – Teses. 3. Sustentabilidade – Teses. 4. Economia – Teses. I. Ribeiro, Rafael Saulo Marques. II. Amaral, Pedro Vasconcelos Maia do. III. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. IV. Título.

CDD: 338.9

Elaborado por Fabiana Santos CRB-6/2530

Biblioteca da FACE/UFMG. – FS/149/2024

UNIVERSIDADE FEDERAL DE MINAS GERAIS

FACULDADE DE CIÊNCIAS ECONÔMICAS

CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

FOLHA DE APROVAÇÃO

"Empirical Essays on Infrastructure and Regional Development: Bringing Highway Investments More Efficient, Inclusive and Sustainable"

VICTOR MEDEIROS

Tese de Doutorado defendida e aprovada, no dia **nove de agosto de dois mil e vinte quatro** , pela Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Economia da Universidade Federal de Minas Gerais constituída pelos seguintes professores:

Carla Cristina Aguilar de Souza

Fundação João Pinheiro

Thiago Costa Monteiro Caldeira

Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa

Ana Maria Hermeto Camilo de Oliveira

CEDEPLAR/FACE/UFMG

Bernardo Palhares Campolina Diniz

CEDEPLAR/FACE/UFMG

Pedro Vasconcelos Maia do Amaral - Coorientador CEDEPLAR/FACE/UFMG

Rafael Saulo Marques Ribeiro - Orientador CEDEPLAR/FACE/UFMG

Belo Horizonte, 09 de agosto de 2024.

Documento assinado eletronicamente por **Rafael Saulo Marques Ribeiro**, **Coordenador(a) de curso de pós-graduação**, em 09/08/2024, às 17:12, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ana Maria Hermeto Camilo de Oliveira**, **Professora do Magistério Superior**, em 10/08/2024, às 11:43, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Pedro Vasconcelos Maia do Amaral**, **Professor do Magistério Superior**, em 12/08/2024, às 10:44, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Carla Cristina Aguilar de Souza**, **Usuário Externo**, em 12/08/2024, às 14:58, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Thiago Costa Monteiro Caldeira**, **Usuário Externo**, em 13/08/2024, às 14:55, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Bernardo Palhares Campolina Diniz**, **Professor do Magistério Superior**, em 19/08/2024, às 15:24, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3453429** e o código CRC **8C7AA5EA**.

Dedico este trabalho à minha saudosa avó, Dona Lourdes, que, sem sombra de dúvida, fez parte e seria a pessoa mais orgulhosa das minhas conquistas.

AGRADECIMENTOS

Iniciei o processo de desenvolvimento dessa tese há pouco mais de cinco anos. Desde lá, é difícil lembrar de se quer uma coisa que não tenha se transformado na minha vida. Não é possível esquecer os momentos mais difíceis de 2019 e 2020, os quais incluíram nada menos que uma longa e desgastante pandemia, um doloroso rebaixamento do clube amado e as diversas e profundas incertezas que a trajetória acadêmica nos colocava em um cenário de poucos recursos e desincentivos à pesquisa no país. Ao mesmo tempo, e mais importante, me emociono em rememorar quantos momentos felizes, quantos aprendizados, quantas conquistas e quanto crescimento eu pude viver durante esse longo, mas prazeroso processo. Escrevo isso com a minha mais profunda gratidão a todas as pessoas que, direta ou indiretamente, me ajudaram e compartilharam comigo esses momentos. Tenho convicção em dizer, sem vocês isso não seria possível. Embora não consiga, aqui, nominar a todos e todas, e peço desculpas por isso, faço alguns agradecimentos especiais.

À minha família. Aos meus pais, Léia e Carlos, que me apoiam desde quando nem consigo me lembrar. A vocês, minha eterna gratidão. À minha saudosa avó Lourdes, que, mesmo em outro plano, continua parte de toda a minha trajetória. À minha irmã, Laís, e meus sobrinhos, por todo o apoio e momentos de alegria insuperáveis. Aos demais familiares que me apoiaram, meu muito obrigado.

A meus orientadores, por todo o auxílio material e intelectual, e pelas grandes contribuições que moldaram esse trabalho. Ao Rafael, por toda a atenção dada desde o mestrado, por tantas portas abertas, as quais nem consigo enumerar, e por todas as discussões e conselhos, no âmbito dessa tese ou não, durante esse processe e que, sem dúvidas, foram essenciais para minha formação acadêmica e profissional. Ao Pedro, por sempre estar disponível e pelas contribuições sempre excepcionais.

A todos os meus amigos e colegas do Cedeplar, que tornaram a pesquisa um processo tão agradável e natural. Agradeço aos camaradas da turma do mestrado de 2017 e do doutorado de 2019 e, em especial, aos grandiosos Pedro, Alexandre, Guilherme, Jonas, Stefan, entre muitos outros e outras. Agradeço, também, ao Cedeplar e todos os seus funcionários por todo o apoio e por tornar esse centro um ambiente tão maravilhoso.

Aos amigos e amigas que, mesmo longe da academia, me apoiaram e tiveram importância enorme nessa conquista. Em especial, agradeço aos meus queridos e queridas Eduardo, Gustavo, Lucas, Vitor, William, Edson, Ítalo, Matheus, Bruno, Tiago, Hugo, João, Laís, Carolina, Vanessa, aos demais amigos da turma da UFV/BH e de Manhumirim.

Agradeço, também, aos colegas da SEINFRA/MG e ICO Consultoria, os quais fizeram parte de um período de grande crescimento profissional e pessoal e que, sem dúvidas, impactou, e muito, nas proposições, no amadurecimento e na qualidade dessa tese.

Agradeço à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) que, por meio de bolsa de doutorado, financiou a realização dessa pesquisa.

Agradeço, finalmente, às professoras Ana Hermeto e Carla Souza e aos professores Bernardo Campolina e Tiago Caldeira pela disponibilidade em participar da banca de defesa desta tese e pelas valiosíssimas contribuições dadas ao trabalho.

ABSTRACT

This dissertation assesses the economic, social, and environmental impacts of highway infrastructure in Brazil. We investigate the relationship between road development and economic outcomes, regional disparities, and sustainability during the Growth Acceleration Program (PAC) period (2007-2018), a case study marked by deep regional heterogeneities in terms of socioeconomic development, infrastructure endowment, and environmental degradation. Chapter 2 develops an original three-step econometric identification strategy to estimate the causal effects of road investments on development outcomes. Using novel granular data on national road investments at the municipal level, we propose several instrumental variables (IVs) to overcome two critical endogeneity issues in infrastructure studies. To correct measurement errors in the road variable, we construct instruments based on the main geographical, environmental, and human-physical infrastructure project costs. To fix the non-random placement of road interventions, we draw different IVs based on global cost minimization methods, historical transportation plans, and the propensity of a municipality to receive road investments. Using our suitable IVs, results suggest that a 1% increase in road investments raises productivity from 0.011% to 0.017%. From these elasticities, we calculate an average Return Rate to Highway Investments (RR) of around 21% in Brazil. Our identification strategy seems accurate under many robustness checks. We use the data and the econometric approach constructed in Chapter 2 in the following empirical exercises. Chapter 3 evaluates the heterogeneous impacts of road investments on productivity, considering efficiency, road specialization, redistribution, and equity goals of public policies. Econometric results point out that the economic impacts of highway investments are higher for poorer and less infrastructure-endowed regions, suggesting that road policies generate a “social bonus” by reducing regional inequalities, and this policy tool might be used for inclusive purposes. From these empirical findings, we augment our return rate measure by slicing the RR into an Economic Component (EC), representing the isolated impact of roads on productivity, and a Social Component (SC), capturing the higher impacts of roads in less developed regions. Chapter 4 measures the environmental costs of highway development. We adapt our three-step identification approach to estimate the impacts of road investments on Greenhouse Gas Emissions (GHG) in Brazilian municipalities. Results show that constructing and enhancing road infrastructure raises GHG emissions. We calculate an average Environmental Discount Rate to Highway Investments (ER) of around 3% in Brazil. We also identify some heterogeneities in the environmental road impacts. The harmful effect of roads on the environment is more pronounced in poorer and more remote areas, especially due to the wider road effect in increasing deforestation. We use the ER as an Environmental Component (EC) and calculate an original Sustainable and Equitable Return Rate to highway investments (SERR) at the regional level. We compute an average SERR of 17%, indicating the high profitability of road investments even considering social and environmental issues. Chapter 5 develops eligibility and prioritization criteria for regionalized road investments considering economic, social, and environmental issues. The eligibility criteria ensure a minimum return level on the investment. The prioritization criteria go beyond the economic aspect conventionally considered in the design of infrastructure projects, ranking higher those regions with the potential to reduce inequalities and mitigate environmental damage through highway investments. Our findings offer novel inputs for policymakers, technicians, financial institutions, and civil society in shaping efficient, equative, and environmentally conscious road policies.

Keywords: transportation infrastructure; economic development; regional inequality; sustainability; highways; roads.

RESUMO

Esta tese avalia os impactos econômicos, sociais e ambientais dos investimentos em infraestrutura rodoviária no Brasil. Investigamos a relação entre o desenvolvimento rodoviário e os crescimento econômico, as disparidades regionais e a sustentabilidade durante o período do Programa de Aceleração do Crescimento (PAC) (2007-2018), um estudo de caso marcado por profundas heterogeneidades regionais em termos de características socioeconômicas, dotação de infraestrutura e degradação ambiental. O Capítulo 2 desenvolve uma nova estratégia de identificação econométrica em três etapas para estimar os efeitos causais dos investimentos em estradas sobre a produtividade. Usando inéditos dados granulares sobre investimentos em estradas nacionais em nível municipal, propomos diversas variáveis instrumentais (IVs) para superar dois problemas críticos de endogeneidade em estudos sobre infraestrutura. Para corrigir erros de medida na variável de infraestrutura rodoviária, construímos instrumentos com base nos principais custos geográficos, ambientais e físicos-humanos em projetos de infraestrutura. Para corrigir problemas de endogeneidade advindos da alocação não aleatória das intervenções rodoviárias, elaboramos diferentes IVs com base em métodos de minimização de custos globais, planos históricos de transportes e a propensão de um município a receber investimentos rodoviários. Usando nossas IVs consideradas adequadas, os resultados sugerem que um aumento de 1% nos investimentos em rodovias aumenta a produtividade entre 0,011% e 0,017%. A partir dessas elasticidades, calculamos uma taxa de retorno econômico média para investimentos em rodovias (RR) de cerca de 21% no Brasil. Nossa estratégia de identificação se mostrou precisa sob vários testes de robustez. Usamos os dados e a abordagem econométrica construída no Capítulo 2 nos exercícios empíricos subsequentes. O Capítulo 3 avalia os impactos heterogêneos dos investimentos rodoviários sobre a produtividade, considerando a eficiência, a especialização rodoviária, a redistribuição e a equidade como objetivos das políticas públicas de transportes. Os resultados econométricos apontam que os impactos econômicos dos investimentos em rodovias são maiores para regiões mais pobres e menos dotadas de infraestrutura, sugerindo que políticas rodoviárias geram um "bônus social" ao reduzir as desigualdades regionais, e que essa ferramenta política pode ser usada para fins inclusivos. A partir dessas constatações empíricas, expandimos nossa medida de taxa de retorno em um Componente Econômico (EC), que representa o impacto isolado das rodovias sobre a produtividade, e um Componente Social (SC), que capta os impactos mais altos das rodovias em regiões menos desenvolvidas. O Capítulo 4 mede os custos ambientais do desenvolvimento das rodovias. Adaptamos nossa abordagem de identificação em três etapas para estimar os impactos dos investimentos em rodovias sobre as emissões de gases de efeito estufa (GEE) nos municípios brasileiros. Os resultados mostram que a construção e a melhoria da infraestrutura rodoviária aumentam as emissões de GEE. Calculamos uma taxa média de desconto ambiental para investimentos em rodovias (ER) de cerca de 3% no Brasil. Também identificamos algumas heterogeneidades nos impactos ambientais das estradas. O efeito prejudicial das rodovias sobre o meio ambiente é mais pronunciado em áreas mais pobres e remotas, especialmente devido ao efeito mais amplo das rodovias no aumento do desmatamento. Usamos a ER como um Componente Ambiental (EC) e calculamos uma original Taxa de Retorno Econômico, Equitativo e Sustentável (TREES) para investimentos em rodovias em nível regional. Mensuramos uma TREES média de 17%, indicando a alta rentabilidade dos investimentos rodoviários no país mesmo considerando questões sociais e ambientais. Por fim, o Capítulo 5 desenvolve critérios de elegibilidade e priorização para investimentos rodoviários regionalizados, considerando questões econômicas, sociais e ambientais. Os critérios de elegibilidade garantem um nível mínimo de retorno sobre o investimento. Os critérios de priorização vão além do aspecto econômico convencionalmente considerado na elaboração de projetos de infraestrutura, classificando como de mais alta prioridade aquelas regiões com potencial para reduzir as desigualdades e mitigar os danos ambientais por meio de investimentos em rodovias. Nossas descobertas oferecem novos insumos para formuladores de políticas, técnicos, instituições financeiras e sociedade civil na elaboração de políticas rodoviárias eficientes, equitativas e ambientalmente corretas.

Palavras-chave: infraestrutura de transporte; desenvolvimento econômico; desigualdade regional; sustentabilidade; rodovias; estradas.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1. INTRODUCTION

Infrastructure investment has been seen as a critical policy tool to foster economic growth and development. Transportation infrastructure improvements reduce trade costs and travel times, raise labor and capital productivity, expand internal and external markets, and intensify agglomeration economies. A massive strand of theoretical and empirical investigations carried out over decades has argued for the presumed positive role of infrastructure investments on the economy (Aschauer, 1989; Baldwin *et al.,* 2001; Behrens *et al.,* 2007, 2008; Bird and Straub, 2020; Duran ton *et al.,* 2014; Faber, 2014; Hirschman, 1958; Krugman, 1993; Ottaviano, 2008; Puga, 1999; Redding and Turner, 2015; Straub, 2011; Zhang *et al.*, 2020), which has inspired governments around the world to use this policy instrument with the promise of boosting economic activity through increasing income and employment.

Although infrastructure investment has the potential to stimulate economic growth and generate prosperity, increased transportation infrastructure has been closely associated with regional disparities (Baum-Snow *et al.,* 2020; Cocci and Mira, 2018; Faber, 2014; Jaworski and Kitchens, 2019; Medeiros and Ribeiro, 2020; Medeiros *et al.*, 2021a, 2022; Zhang and Ji, 2019). Infrastructure interventions are spatial by nature, serving specific geographic areas aimed at linking important economic centers to other economic zones, highly populated regions, strategic ports, and so forth, developing lagging regions or attending political purposes. On the one hand, *win-win* situations wherein economic growth is achieved with reduced inequality and poverty alleviation might emerge. In these cases, infrastructure projects have the potential to help bridging the gap between different regions by making it easier for businesses to move to less developed areas with lower operating costs and better market access. Additionally, constructing infrastructure might facilitate the spread of knowledge, leveling the playing field between regions regarding production costs. On the other hand, infrastructure development might bring national economic growth, but this wealth may reach some areas at the expense of others. It occurs when enhancing connectivity between a remote region and a highly developed area amplifies the market benefits experienced by the latter, prompting businesses to relocate there. Finally, there is the case in which infrastructure investment promotes a regional reorganization of economic activity from one area to another, with null net impacts on the economy. In this complex context, the spatial placement of infrastructure is critical in shaping the regional economy, and prioritizing regions wherein high economic returns are accompanied by decreased inequality has a pivotal role in maximizing the broader benefits of transportation investments.

In addition, infrastructure improvements are deeply related to environmental damage (Asher, 2020; Churchill *et al.*, 2021; Emodi *et al.*, 2022; Lin and Chen, 2020; Santos, 2017; Yao *et al.*, 2023). Road construction and improvement increase GHG emissions during the building and maintenance phases through the direct use of materials and equipment. Once built, the new infrastructure enhances regional accessibility and mobility, impacting GHG emissions through increased transportation demand. Moreover, transportation infrastructure development might enlarge deforestation, water pollution, ecological footprint, and other environmental outcomes. Conversely, road improvement may reduce pollution by decreasing travel time and distance, promoting agglomeration economies, and improving energy efficiency. Besides, the net effect of road investments on the environment will be affected by a range of regional features such as agglomeration economies, income levels, population scale and density, technology intensity, institutional quality, legal environmental protection framework, and so forth. Then, road investments might heterogeneously impact the environment, and policymakers should consider it when allocating road interventions across regions.

Despite its evident role in the transportation sector, social and environmental benefits (costs) have been broadly overlooked when measuring the returns to infrastructure investments. Most studies have focused on the economic issues of road investments, mainly by evaluating its impacts on income, travel times, and reducing transportation costs (Alam *et al.*, 2022; Quadros and Nassi, 2015; Laird and Venables, 2017; Welde and Tveter, 2022). These results have been extensively applied, for instance, in Cost-Benefit Analysis (CBA), Value for Money (VfM), and several other infrastructure project evaluation methods worldwide, influencing national, regional, and local governments to elect road interventions primarily based on the economic criteria, neglecting, or downplaying the social and environmental components of road investments.

This dissertation sheds light on those issues by estimating the economic returns of highway investments conditioned by social and environmental factors in Brazil during the Growth Acceleration Program (PAC) period. The Brazilian case is a unique empirical setting to evaluate the impacts of transportation infrastructure and how they affect regional inequalities and the environment. First, road investments have been underperforming for decades since the 1980s, resulting in poor infrastructure stock, quality, and access. To overcome these shortcomings and stimulate economic growth, the Brazilian Federal Government launched the PAC in 2007 with the promise to develop the infrastructure sector, duplicating the levels of highway investments in comparison with the previous decade (Medeiros *et al.*, 2021b). Then, the PAC constitutes a rare episode of a major expansion of transportation infrastructure in a developing country. Second, Brazil is one of the most unequal countries in the world. This condition is exacerbated by huge spatial heterogeneities in infrastructure endowment, productivity levels, socio-environmental characteristics, and so forth (Medeiros *et al.*, 2021a, 2022). Third, Brazil has faced several environmental challenges over the past decades. Unlike most developed countries and many developing economies, the central contributors to polluting gas emissions are land use change and agriculture sectors, which are deeply associated with deforestation and economic dynamics in the country. Those regional complexities are worthwhile inputs and make Brazil an outstanding case study to evaluate the economic, social, and environmental effects of transportation infrastructure investments wherein deep regional heterogeneities are expected to exist.

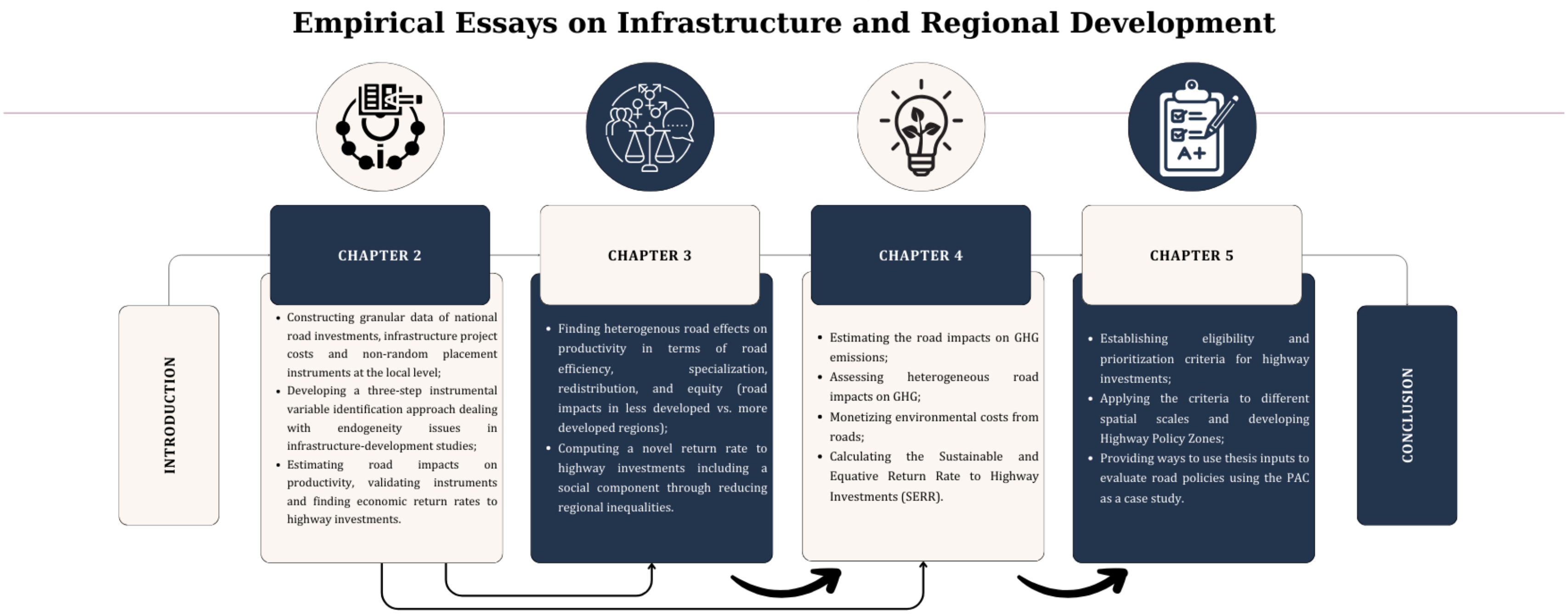
Moreover, in August 2023, the Brazilian Federal Government instituted the “New” PAC, with predicted investment values of around R$ 1.7 trillion in several infrastructure sectors. The main goals of the new (and third) program are similar in comparison to PAC 1 (2007-2010) and PAC 2 (2011-2014)[1], being developing the precarious national infrastructure by augmenting public investments and attracting private resources to the sector. In addition, to the best of our knowledge, this marks the first instance in Brazilian history in which a comprehensive national infrastructure initiative incorporates explicit socio-environmental strategies. As one of the main mechanisms to foster inclusive and environmentally friendly practices in the infrastructure sector, the Brazilian Federal Government prioritizes and facilitates the availability of funds to projects with socio-environmental characteristics. In the transportation sector, the third PAC introduces the "Efficient and Sustainable Transport" pillar, allocating approximately R$ 349.1 billion in investments towards various transportation infrastructures, roads being the most relevant of them. Additionally, the transport program encompasses numerous institutional initiatives to foster social and environmentally conscious road infrastructures to accelerate an ecological and inclusive transition. For instance, the program expands financial resources and facilitates debentures issuing for projects with social, climate and environmental benefits. Furthermore, the new PAC encourages ecological transition by issuing sustainable sovereign bonds, expanding resources to the Climate Fund (*Fundo Clima*), endorsing low-carbon transportation options like hybrid and electric vehicles, and promoting decarbonization and utilizing sustainable materials in the construction sector.

While those policy devices are critical to the Brazilian economic, inclusive, and sustainable development, a precise regionalized measure of highway investments' social and environmental costs (or benefits) is lacking. In this context, an evaluation of the “old” PACs – in which emphatic social and especially environmental initiatives related to the road sector were most part absent –is critical to provide evidence on the social and environmental gains (losses) from road investments, maximizing its economic returns while reducing inequalities and respecting environmental preservation and recovery. Therefore, a novel measure of inclusive and sustainable return rate to highway investments might represent a key input to policymakers, technicians, financial institutions, and civil society in planning, designing, financing, and evaluating current and future road policies. This doctoral thesis contributes to the specialized literature on infrastructure and regional development following this path. Figure 1.1 outlines the arrangement and the rationale behind this dissertation.

---

[1] Whilst the PAC 2 was officially carried out between 2011 and 2014, several unfinished and delayed infrastructure buildings continued for several years at least until 2018. Even today, we can observe some remaining infrastructure works from the PAC 2.

Figure 1.1. Thesis structure and rationality



Source: authors' elaboration.

In Chapter 2, we assess the (economic) road impacts on productivity. We combine PAC intervention information with the georeferenced National Highway System data to create a novel and granular dataset of federal road investments at the municipal level in Brazil between 2007 and 2018. Next, we develop an original econometric strategy dealing with the main endogeneity issues in road-development studies. First, the non-random placement of road policies poses challenges in determining their causal effects on economic activity. Planners may allocate resources to regions where infrastructure promises higher returns to induce national economic growth. Conversely, policymakers might target less developed and more remote areas to foster a more balanced economic development among regions. If it occurs, conventional Ordinary Least Squares (OLS) estimates are likely upward biased in the former context while downward biased in the latter. Second, highway investment often involves measurement errors from inefficiencies and corruption during the project design, construction, and operation phases (Kenny, 2009; Straub, 2011). Consequently, road-related variables are prone to be overstated, with some areas featuring inflated investment values per kilometer of road. As a result, measurement errors may obscure the true impact of roads on economic growth and development, leading to an underestimation of their profitability. To fix those issues, we construct a three-step instrumental variable (IV) identification strategy. The first *ex-ante* step consists of designing a study that minimizes endogeneity concerns by evaluating the impacts of a national road program on local productivity, then alleviating broad endogeneity issues coming from the non-random placement of roads by the Federal Government. In the second step, we instrumentalize our road variable using some of the main geographical, environmental, and human physical costs in infrastructure projects to correct measurement error bias. In the third step, we combine our second-step cost-related IVs with several road allocation IVs to correct for omitted variable bias coming from the non-random placement nature of highway policies. From this, we identify relevant biases from measurement error and omitted variables. Our preferred estimates point out consistent road elasticities between 0.011 and 0.017, implying a non-biased (economic) return rate to highway investments of around 21.3% in Brazil. We use the data and the empirical strategy built in Chapter 2 in the remaining chapters of this thesis.

In Chapter 3, we include a social component into the return rate to highway investments by evaluating regional inequality issues. More specifically, we assess the role of transportation infrastructure policies in raising productivity by considering spatial heterogeneities in terms of efficiency, road specialization, redistribution, and equity. On the economic side, the efficiency goal is tied to the intention of road policies to maximize economic benefits by placing infrastructure in regions with greater growth potential. Additionally, this profitability could be further boosted by improving roads in areas more specialized in the transportation sector. On the inclusion side, redistribution is the policy objective utilizing road interventions to promote balanced economic growth across regions by focusing on poorer areas. Similarly, equity entails investing in places with constrained infrastructure endowment, thus leveling the playing field by regions. By evaluating the road impacts on productivity varying the levels of efficiency, road specialization, redistribution, and equity, we calculate an efficient-specialized and redistributive-equative return rate to highway investments in Brazil.

We identify three main findings in Chapter 3. First, highway investments prove to be highly beneficial in Brazil, and areas more reliant on road infrastructure benefit more from it. Second, the road impacts on productivity are more pronounced in less developed municipalities, presenting poor infrastructure, lower productivity levels and relying more on roads. Third, the profitability of road investments appears to suffer significantly when targeting highly efficient and developed areas, likely due to substantial infrastructure project costs and inefficiencies. In other words, focusing solely on highly efficient and affluent regions based on expected economic returns may overlook their inefficiencies, thus overestimating the economic impact of road investments. Based on our econometric results, we estimate an average (economic) return rate to highway investments during the PAC to be around 20%. When considering the road features heterogeneities, this return rate drops to 11.7%, suggesting that the program excessively targeted more prosperous areas wherein road investments did not play a critical role in raising productivity. In other words, the (economic and social) PAC return rate could have been higher if the program had targeted some *win*-w*in* places wherein economic returns were expected to hold with redistribution and equity, which does not seem to have been the case.

In Chapter 4, we extend our (economic and social) return rate to highway investments by incorporating an environmental component. To do this, we evaluate the impact of highway investment on GHG emissions growth in Brazilian municipalities between 2007 and 2018. From these estimates, we calculate a carbon dioxide equivalent emissions return (discount) rate (ERR), which allows us to compute the Sustainable Return Rate (SRR) and the Sustainable and Equitable Return Rate (SERR) to highway investments to several Brazilian localities. Then, our preferred SERR includes economic, social, and environmental road investment benefits (costs) by combining the road impacts on productivity, varying those impacts by social road features such as redistribution and equity, and discounting the highway impacts on GHG emissions.

We obtain three main results in Chapter 4. First, we find that a 1% increase in road investments raises GHG emissions by 0.025%. Second, we calculate an average GHG emissions (discount) return rate to highway investment (ERR) of 3.0% by blending our estimated elasticities and GHG data with Social Cost of Carbon (SCC) measures, demonstrating a harmful environmental impact of roads. This implies an average SERR of around 17%, indicating a widespread need to develop the Brazilian transportation sector, even considering its environmental and social components. Third, we find critical regional heterogeneities in our ERR, SRR, and SERR. In general, the environmental damage from roads is more pronounced in less populated and poorer localities, which coincides with some critical areas in the Brazilian Amazon. In those cases, improving the project governance is a vital issue, including the coordination between national, regional, and local transportation and environmental institutions.

Chapter 5 introduces an innovative empirical methodology to delineate priority regions for highway investments in the Brazilian context. Unlike preceding research primarily emphasizing economic factors (Fernald, 1999; Li *et al.*, 2017; Medeiros *et al.*, 2021b; Wang *et al.*,2020), our approach integrates economic, social, and environmental considerations. By doing so, we seek to identify areas where these three dimensions intersect, recognizing the potential for infrastructure policies to foster inclusive and sustainable economic development.

Achieving optimal outcomes across economic, social, and environmental domains is often challenging. Infrastructure investment decisions frequently prioritize anticipated economic gains while overlooking socio-environmental implications. Consequently, regions with higher expected profitability tend to attract more public and private resources, potentially exacerbating regional disparities and environmental degradation. Conversely, underdeveloped and geographically remote areas might be left behind due to their smaller economic returns and higher construction costs, particularly during periods of fiscal constraint. Given these regional dynamics, a clear prioritization standard for investment allocation becomes vital for maximizing economic returns, addressing regional disparities, and mitigating environmental risks associated with road development.

Then, in Chapter 5, we reassess the findings obtained in Chapters 2, 3, and 4, estimating the Sustainable and Equitable Return Rates to Highway Investments (SERR). From these three empirical exercises, we apply economic, social, and environmental criteria and clustering methods to identify potential prioritization areas for road interventions. We propose criteria for determining eligibility and prioritization of road policies at the regional level. The eligibility standards ensure that projects offer attractive economic prospects while considering social and environmental considerations. Meanwhile, the prioritization criteria categorize regions based on their potential economic profitability, aptitude to mitigate regional disparities, and capacity to minimize environmental harm from road development. In this step, we identify priority regions around the country, suggesting that road investments might be a critical policy tool to foster inclusive and sustainable economic growth. Subsequently, our analysis expands to encompass broader spatial contexts. We employ clustering techniques to delineate highway policy zones, minimizing regional differences in economic, social, and environmental road-related issues. This endeavor yields original insights valuable for formulating and assessing national road policies, particularly those traversing vast geographical expanses and diverse regions.

Finally, we conduct an *ex-post* evaluation of the "old" PACs, suggesting avenues for applying our return rates and prioritization criteria. This evaluation enables us to state that the PAC demonstrated economic efficacy, as it directed investments towards regions yielding economic returns surpassing the conventional cut-off rates. However, hindered social and environmental repercussions somewhat offset these economic gains. Therefore, greater returns (around 31%) on road investments could have been attained by directing resources towards our identified priority regions, thereby fostering *win-win* scenarios characterized by a more inclusive and sustainable economic development.

These essays contribute to the literature strand by examining the impacts of transportation infrastructure on regional development in several ways. First, we provide novel granular data on national highway investments at the local level and an original three-step IV identification strategy to overcome critical endogeneity issues in infrastructure development studies. In addition, we propose some unique instruments for road measures based on actual infrastructure project costs and the non-random placement nature of transport policies. We are convinced that our empirical scheme can be applied, adapted, or extended by several researchers worldwide, contributing to the broad empirical literature evaluating causal road impacts on regional development, especially to those investigations using reduced-form approaches. Second, we extend the conventional economic return rates to highway investments by including social and environmental aspects, and then computing our Sustainable and Equitable Return Rate to Highway Investments (SERR). The social component monetizes the gains from investing in roads in *win-win* (economically profitable and reducing inequalities) regions marked by low productivity and infrastructure endowment, and high efficiency and road specialization. The environmental component monetarily discounts the harmful road impacts on GHG emissions. Then, our SERR makes different aspects of highway investments compatible in an easy-to-interpret measure, allowing society and various public and private players to assess the economic profitability of roads, considering inclusion and sustainability matters. To the best of our knowledge, this is the first regionalized return rate to road investments matching economic, social, and environmental issues, thus providing an important contribution to the empirical literature evaluating the infrastructure impacts on development. Third, we find road impacts on the economy and the environment to be heterogeneous. On the one hand, highway investments are more effective in regions characterized by lower productivity levels and infrastructure endowment. On the other hand, road interventions harm the environment more in poorer, less populated, and more geographically isolated areas, especially through an indirect effect on GHG emissions in the land use change sector. From these results, we contribute to several studies identifying heterogeneous road impacts on development outcomes. Finally, we furnish original eligibility and prioritization criteria to categorize regions in terms of economic returns, reducing inequalities and minimizing environmental damages. This classification allows policymakers and the society to evaluate past, current, and future road policies, providing novel inputs to bring road investments more efficient, inclusive, and sustainable. Then, we collaborate with several investigations evaluating road policy efficiency, efficacy, and effectiveness, especially those concerned with regional disparities and the environment. A more detailed description of this dissertation's main and marginal contributions to the specialized literature can be seen in each chapter.

2. HIGHWAY INFRASTRUCTURE AND ECONOMIC DEVELOPMENT: measuring causal impacts of infrastructure investments using a three-step instrumental variable identification strategy

Abstract

This paper provides an original three-step identification strategy using instrumental variables to evaluate the causal impact of highway investments on the local economy. First, we construct a novel national highway dataset at the municipal level in Brazil using the Growth Acceleration Program (PAC) (2007-2018) as a case study. Second, we rely on some of the main infrastructure project costs to propose several cost-related instruments to correct measurement errors in the road variables. Third, we circumvent the omitted variable bias from the non-random placement of roads by building instruments based on global cost minimization methods, historical plans, and the propensity of a municipality to receive highway interventions. Our identification strategy allows us to identify relevant biases from measurement errors and omitted variables. Our preferred estimates point out a reliable road elasticity in the range of 0.011 to 0.017. From this, we calculate a non-biased return rate to highway infrastructure of 21.3% in Brazil, proving the high rentability of those investments in the developing world context.

Keywords: highway infrastructure; regional development; endogeneity; instrumental variables.

2.1. Introduction

The economic effects of transportation infrastructure have been explored by various studies (Anas, 2020; Aschauer, 1989; Baum-Snow *et al.*, 2017; Baum-Snow *et al.,* 2020; Chandra and Thompson, 2000; Donaldson and Hornbeck, 2016; Duranton and Turner, 2012, 2014; Faber, 2014; Farhadi, 2015; Foster *et al.*, 2023a, 2023b; Michael, 2008; Straub, 2011). More interestingly, a substantial part of this literature has provided relevant findings on highway investments' role in local economic activity. Generally, the findings indicate a positive and direct association between road investments and various outcomes, such as economic growth, productivity, employment, and poverty alleviation, among others (Allen and Arkolakis, 2014; Arbués *et al.*, 2015; Bird and Straub, 2020; Ghani, Goswami, and Kerr, 2014; Holl, 2016; Huang and Xiong, 2018; Li *et al.*, 2017; Percoco, 2015; Wang, Wu, and Feng, 2020).

Despite the long-standing development of this literature, endogeneity bias remains a persistent empirical challenge (Calderón and Servén, 2014; Redding and Turner, 2015; Roberts *et al.*, 2019; Straub, 2011). First, the non-random allocation nature of road policies makes it hard to identify causal impacts. Planners may allocate more resources to regions where infrastructure yields higher returns to foster national economic growth. On the other hand, policymakers may target less developed and more remote regions to promote more equitable economic development across regions. If it occurs, conventional Ordinary Least Squares (OLS) estimates are likely upward biased in the former context while downward biased in the latter. In addition, highway investment measures often embed measurement errors due to inefficiencies and corruption in the infrastructure project design phase and its building and operation stages (Kenny, 2009). Then, road variables are likely to be inflated because, for many places, we usually observe an overpriced amount of investment *per* kilometer of road. Consequently, the measurement error may obscure the real impact of roads on economic growth and development, resulting in an underestimation of their profitability.

To solve the endogeneity issue, most empirical studies have used reduced-form estimations under several different instrumental variables (IV) as sources of quasi-random variation in the observed infrastructure (Foster *et al.*, 2023a, 2023b; Redding and Turner, 2015; Roberts *et al.*, 2019). Investigations have proposed instruments based on either planned routes (Baum-Snow, 2007; Bird and Straub, 2020; Duranton and Turner, 2012; Duranton *et al.*, 2014; Hsu and Zhang, 2014; Michaels, 2008; Herzog, 2021; Rokickia and Stępniak, 2018), or historical routes (Adler *et al.*, 2020; Baum-Snow *et al.*, 2017; Baum-Snow *et al.*, 2020; Duranton and Turner, 2012; Duranton *et al.*, 2014; Garcia-López *et al.*, 2015; Holl, 2012; Holl, 2016; Hsu and Zhang, 2012; Lee, 2021; Martın-Barroso, Nunez-Serrano and Velazquez, 2015; Martincus *et al.*, 2017; Percoco, 2015; Rokickia and Stępniak, 2018; Zhang, Hu and Lin, 2020), or infrastructure project costs (Holl, 2012; Martin-Barroso, Nunez-Serrano, and Velazquez, 2015; Lu *et al.*, 2022; Medeiros *et al.*, 2021a; Medeiros *et al.*, 2021b; Zhang, Hu, and Lin, 2020) or hypothetical road networks built on a global minimization path intended to connect important localities (hubs) (Faber, 2014; Ghani, Goswami, and Kerr, 2014; Huang and Xiong, 2018; Yang, 2018; Xu and Feng, 2022). A relevant part of those studies has combined some IV with the inconsequential unit approach pioneered by Chandra and Thompson (2000). We also suggest Redding and Turner (2015) and Foster *et al.* (2023a, 2023b) as reference papers.

While this strand of literature has developed interesting and robust identification strategies for measuring the causal highway impact on several outcomes, some open points remain. First, instruments are hard to find in practice, and a highly replicable IV approach is needed. Second, researchers have avoided monetary road variables due to measurement errors (Calderón and Servén, 2014; Straub, 2011). With the increase in public administration transparency worldwide in the last decades, investment and expense flow data might be an important source of information on the efficacy, efficiency, and effectiveness of infrastructure policies. In addition, highway investment variables constitute a direct way of measuring infrastructure profitability and providing an easy-to-understand indicator for planners and society (Fernald, 1999; Li *et al.*, 2017; Wang, Wu, and Feng, 2020). Therefore, a reliable identification strategy dealing with measurement errors in highway measures is critical. Third, monetary variables have the advantage of capturing both road provision and quality. Most empirical studies have used measures of physical provision , and empirical approaches to adapt a multitype road intervention setting are still lacking. Fourth, a more careful look into the study design is demanding. For instance, recent investigations have evaluated the impact of national road programs on local outcomes, an approach that considerably reduces endogeneity concerns and has provided reliable results (Bird and Straub, 2020; Faber, 2014; Ghani, Goswami, and Kerr, 2014; Herzog, 2021).

We propose an original three-step identification strategy using IVs to overcome the mentioned issues and estimate the causal impact of highway investments on local outcomes. The first *ex-ante* step consists of designing a study that minimizes endogeneity concerns. For this, we use data from a national road policy, the Growth Acceleration Program (PAC)[2], launched by the Brazilian Federal Government in 2007. We georeferenced the PAC data to construct a novel granular national highway investment data at the municipal level.

The PAC has some characteristics that make it an interesting case study. First, highway investments were duplicated during the PAC (2007-2018) compared to the previous ten years (Medeiros *et al.*, 2021a). Second, the PAC coincided (up to 2015) with a relatively high economic growth period in Brazil (Nassif *et al.*, 2015). Third, the program faced severe criticism for its inefficiency and poor budget management, as it exhibited multiple construction delays and required much more significant investments than the original estimates (Amann *et al.*, 2016; Raiser *et al.*, 2017). In this sense, this developing economy case study seems appropriate for evaluating the infrastructure-development nexus wherein measurement error is highly expected.

Then, we proceed to our second and third steps to evaluate the causal impact of national highway investments on municipal GDP *per capita* growth in Brazil between 2007 and 2018. In the second step, we instrumentalize our road variable using some of the main

[2] The PAC included several infrastructure sectors. In this study, we focus on road transportation.

geographical, environmental, and human physical costs of infrastructure projects to correct measurement error bias. We rely on the empirical literature and Brazilian real infrastructure projects to develop replicable and reliable cost-related IVs. Conditional on controls, the second-step econometric estimations allow us to identify suitable (strong and exogenous) cost-related IVs.

In the third step, we combine our second-step cost-related IVs with several road allocation IVs to correct for omitted variable bias coming from the non-random placement nature of highway policies. We draw upon the specialized literature and adapt it to the Brazilian case to generate IVs treating investment in new and existing roads. The first set of instruments is intended to treat economically and politically biased road policies for newly connected municipalities. In this case, we use the Least Cost Path - Minimum Spanning Tree (LCP-MST) method (Faber, 2014) combined with cost-related measures to create hypothetical minimized global cost network instruments. In addition, we utilize the Brasília Plan (Bird and Straub, 2020) to generate additional and more specific instruments to treat political bias in road placement. For already connected municipalities in the starting period, we build up a unique instrument using road traffic data to avoid endogeneity from municipalities highly prone to receive road interventions.

Our results show that our empirical strategy is suitable for identifying causal road impacts on the local economy. Firstly, the second-step estimates allow us to correct (or mitigate) measurement error and to determine expected downward biased OLS elasticities. Secondly, third step regressions are critical to fix non-random placement of roads bias even after fixing measurement error. In this case, our findings suggest that the PAC prioritized economic development over regional balance, as its actions favored more prosperous areas, thus upward biasing "free from measurement error" second-step estimates. We find a solid third-step road elasticity ranging from 0.011 to 0.017, implying a return rate to highway investment of around 21.3% in Brazil. From that, we can infer that Brazil would need to invest 2.5 more in road infrastructure to achieve a suitable road stock of 16% of the national GDP, proving the high rentability of highway investments in the developing economy scenario. Results remain unchanged under several robustness checks.

Our contributions are manifold. First, we develop a novel three-step IV identification strategy to correct measurement errors and non-random road allocation bias. We believe that our steps and instruments have a high degree of replicability. Thus, we contribute to an extensive strand of the literature using reduced-form equations to estimate the causal impacts of highway investments on the economy (Bird and Straub, 2020; Chandra and Thompson, 2000; Faber, 2014; Herzog, 2021; Michaels, 2008), as well as to the overall empirical literature on infrastructure.

Second, we construct unpublished granular highway investment data at the municipal level. This data is particularly relevant as geographically detailed data on infrastructure investment is relatively scarce (Brooks and Liscow, 2019). In this sense, we contribute to the broad infrastructure-development literature that is interested in measuring highway investments and their impact on the economy at the local level.

Third, we propose several original and replicable cost-related instruments in our second step. We rely on real infrastructure projects to propose geographical, environmental, and human physical measures to represent road costs. To the best of our knowledge, our study is the first to jointly use those cost types as IVs in the context of measurement error bias. In addition, our study tests some of our proposed measures for the first time. We complement a range of works using cost-related IVs for transport measures (Holl, 2012; Martin-Barroso, Nunez-Serrano, and Velazquez, 2015; Lu *et al.*, 2022; Medeiros *et al.*, 2021a, 2021b; Zhang, Hu, and Lin, 2020).

Fourth, we contribute to the empirical literature using LCP-MST instruments for road variables in two ways (Faber, 2014; Ghani, Goswami, and Kerr, 2014; Huang and Xiong, 2018; Yang, 2018; Xu and Feng, 2022). First, we construct a cost index based on our preferred infrastructure project cost variables and include it in the minimization process in the LCP-MST method. By doing so, we improve the instruments based on the LCP-MST by including different kinds of road costs. Second, we propose a more replicable way of establishing the hubs the LCP-MST procedure is connecting. We do this by identifying starting and ending points of roads constructed or improved by the PAC, which we believe is a reasonable approach when a clear policy identification of targeted cities is absent. Similarly, we complement the study by Bird and Straub (2020) by proposing an extension of their Brasília Plan instrument.

Fifth, we attempt to empirically deal with the multitype road intervention setting (building, paving, enhancements, and duplications) of our data. Our econometric specification differentiates municipalities already connected by federal roads in 2006 from those not linked. Then, we use LCP-MST and Brasília Plan instruments for non-connected municipalities while we create a novel instrument for already connected ones. This original instrument, called "*potential road intervention areas*" IV, is calculated by identifying critical points using road traffic data, a measure we believe has not been tested in past studies.

The paper is structured as follows. Section 2 is a literature review. Section 3 presents the rationality behind our three-step IV identification approach. Section 4 details the different sources of data used in the paper. Section 5 introduces the empirical model. Section 6 presents the main results and robustness checks. Section 7 discusses our results considering the return rate to highway investments. Section 8 concludes. Further discussion and robustness checks can be found in the Supplementary Materials, identified as Appendices A to J.

2.2. Related literature

2.2.1. Road infrastructure and regional or local economic development: empirical related literature

An extensive strand of literature has examined the relationship between road infrastructure and several economic factors such as output growth (Barzin *et al*, 2018; Baum-Snow *et al.*, 2017; Baum-Snow *et al.*, 2020; Bird and Straub, 2020; Ke and Yan, 2021; Rokickia and Stępniak, 2018; Zhang, Hu and Lin, 2020), productivity (Ghani, Goswami and Kerr, 2014; Fahardi, 2015; Holl, 2016; Huang and Xiong, 2018; Li *et al.*, 2017; Martın-

Barroso, Nunez-Serrano and Velazquez, 2015; Xu and Feng, 2022; Yang, 2018; Zhang e Ji, 2019), trade (Coşar and Demir, 2022; Duranton *et al.*, 2014; Martincus *et al.*, 2017), population (Adler *et al.*, 2020; Baum-Snow, 2007; Baum-Snow *et al.*, 2017; Baum-Snow *et al.*, 2020; Bird and Straub, 2020; Duranton and Turner, 2012; Faber, 2014; Garcia-López *et al.*, 2015; Gertler *et al.*, 2019, 2022; Jaworskiy and Kitchensz, 2019; Meijers *et al.*, 2012; Percoco, 2015) and structural transformation (Albalate and Fageda, 2016; Asher and Novosad, 2020; Yang, 2018). Since Aschauer (1989), most studies have shown a significant relationship between infrastructure and economic-related outcomes. However, results vary critically according to the investigation context and the identification strategy used, which is associated with endogeneity issues and with whether the study uses aggregate data at the country, regional, or local level.

In this paper, we focus on regional and local level studies. Infrastructure is spatial by nature (Ottaviano, 2008; Straub, 2008, 2011), and a more geographically disaggregated view of the theme can clarify some transmission channels. As transport infrastructure buildings serve a limited geographic zone, the results from those policies might expand economic growth in some regions and sectors at the expense of others[3]. This is likely why works using regional and local level data provide more heterogeneous results on the role of road investment on economic activity (Foster *et al.*, 2023a, 2023b; Redding and Turner, 2015; Roberts *et al.*, 2019).

Then, we turn to the main empirical issue in the infrastructure-economic development literature: endogeneity bias. First, measurement error bias often occurs, especially when using monetary infrastructure variables. Highway investments take time to mature and suffer from inefficiencies and delays, particularly in the developing world context (Calderón and Servén, 2014; Kenny, 2009; Straub, 2011). Second, the placement of roads is not random, as policymakers may target places expected to grow more or promote economic prosperity in underdeveloped regions (Redding and Rossi-Hansberg, 2017; Redding and Turner, 2015). In this sense, we can expect endogeneity issues from omitted variables and reverse causality.

Roberts *et al.* (2019) study a wide range of papers estimating the effects of transportation infrastructure on several outcome variables. Around 63.5% used identification strategies to address endogeneity concerns, of which 78% used reduced form estimations, and 52.6% used IVs. Therefore, most papers rely on IVs, and studying this identification approach in detail is critical.

In the context of IV studies, the main bottleneck is the identification of suitable (strong and exogenous) instruments for highway measurement. In successful applications, this identification strategy has been quite reliable in identifying the causal impacts of road investments on local outcomes (Baum-Snow, 2007; Duranton and Turner, 2012; Faber, 2014; Holl, 2012). Nonetheless, essential efforts are still needed, especially in the developing economy context where data is rough and measurement errors are expected to exist in road variables. This issue is accentuated by using monetary variables

[3] An extensive literature on Economic Geography has examined the infrastructure issue. For a more detailed literature review, we recommend Straub (2008) and Ottaviano (2008).

as infrastructure investments or expense flows if those variables might embed inefficiencies and mask the true relation between transport infrastructure and economic activity. The following section covers the main IV approaches used by the specialized literature.

2.2.2. Main IV approaches in studies on the impacts of road infrastructure on regional or local development

Most empirical literature on infrastructure and regional or local development has used IVs to identify causal road impacts. These studies can be classified into a few main approaches based on the IV rationale and measurement form. We draw upon the seminal paper by Redding and Turner (2015), and then we update this literature review and propose some advances. To this end, we also rely on Foster *et al.* (2023a, 2023b), which reviewed a massive number of empirical studies on infrastructure and development, including the road sector.

Redding and Turner (2015) identified three main IV strategies. The first one, the planned route approach, is an IV strategy relying on planning maps and old documents as a source of quasi-random variation for the observed infrastructure (Baum-Snow, 2007; Bird and Straub, 2020; Duranton and Turner, 2012; Duranton *et al.*, 2014; Frye, 2016; Hsu and Zhang, 2014; Michaels, 2008; Sheard, 2014; Herzog, 2021; Rokickia and Stępniak, 2018). The rationale behind these IVs is that the planned routes were created to serve purposes quite different from those from modern infrastructure development we are trying to test (for example, impacts on modern GDP, population, or employment growth). Planned routes used in this approach are, for instance, the 1947 USA National Interstate Highway Plan and the "Pershing plan" for US data, the Brasilia Road Plan for Brazilian data, Japan's 1987 National Expressway Network Plan applied to Japan, and the project for a motorway network by E. Buszma dating from 1945 and the Resolution of the Automotive Council for the Council of Ministers plan from 1963 for the Polish case. A similar strategy was used by Sheard (2014) applied to the airport sector, using the USA 1944 National Airport Plan as IV.

The second approach, the historical route instrumental variable approach, relies on very old transportation routes as a source of quasi-random variation for observed infrastructure (Adler *et al.*, 2020; Baum-Snow *et al.*, 2017; Baum-Snow *et al.*, 2020; Duranton and Turner, 2012; Duranton *et al.*, 2014; Garcia-López *et al.*, 2015; Holl, 2012; Holl, 2016; Hsu and Zhang, 2012; Lee, 2021; Martın-Barroso, Nunez-Serrano and Velazquez, 2015; Martincus *et al.*, 2017; Percoco, 2015; Rokickia and Stępniak, 2018; Zhang, Hu and Lin, 2020). The validity of this kind of instrument requires that, conditional on controls, factors that do not directly affect economic activity in the localities of interest at the period of study (mainly at the end of the twentieth century) determine the configuration of these historical networks. In other words, the validity of this identification strategy depends critically on the fact that the historical routes served different aims (for instance, moving agricultural goods to local markets, achieving military, administrative, and commercial goals, and so forth) than they have today (for example, to foster economic growth and employment). Several distinct IVs were proposed in this setting, including the 1898 railroad routes and the routes of major expeditions of exploration between 1528–1850 for US data, a market access measure based on railroad network in 1870 for European regions data, the network of Roman roads for Italy data, the old Roman roads, the 1760 Bourbon roads, the 1760 Spanish postal route network and an accessibility to final markets in 1857 in the Spanish case, railway connections in 1952 for Polish data, locations of rail and tram stations in 1897, 1930, and 1960 for South Korea data, road and railroad networks in 1962 and postal routes in 1936 for China, and the Pre-Columbian Inca road network for the Peruvian case.

The third one, the inconsequential unit approach, relies on choosing a sample that is inconsequential to road allocation in the sense that unobservable attributes do not affect the placement of infrastructure (Chandra and Thompson, 2000). Sometimes, this strategy is combined with some planned, historical, or cost-related IV (Banerjee *et al.*, 2012; Bird and Straub, 2020; Faber, 2015; Herzog, 2021; Percoco, 2015). In this approach, routes that sought to connect large centers cross (inconsequentially) small units (as a small city or municipality) across the pathway. Then, we expect that unobserved characteristics of these small units, which are just between targeted hubs, are independent of political and economic reasons.

Another strand of literature has utilized a Least Cost Path - Minimum Spanning Tree (LCP-MST) IV (Faber, 2015; Frye, 2016; Ghani, Goswami, and Kerr, 2014; Huang and Xiong, 2018; Yang, 2018; Xu and Feng, 2022). This identification strategy tries to answer the question of which routes planners would have been likely to build if the sole policy objective had been to connect all targeted city nodes on a single continuous network subject to global construction cost minimization. To this end, an LCP-MST network connecting large cities (hubs) is generated based on a global cost minimization process. The identifying assumption is that this hypothetical highway network should affect city outcomes and the spatial allocation of industries only through the actual highway network, conditional on controls. The LCP-MST approach has been used especially for Chinese data, with some applications for the USA and India.

A fifth approach uses cost-related environmental and geographical IVs as a source of quasi-random variation for highway investments (Holl, 2012; Martin-Barroso, Nunez-Serrano, and Velazquez, 2015; Lu *et al.*, 2022; Medeiros *et al.*, 2021a; Medeiros *et al.*, 2021b; Zhang, Hu, and Lin, 2020). The rationale behind this identification strategy is that geographical costs directly affect road construction and maintenance. However, they are exogenous factors in that they do not directly relate to modern GDP, employment, wages, or productivity growth. These IVs are mainly measured as slope, terrain ruggedness, altitude, or elevation, and they have been used for Chinese, Spanish, and Brazilian data. The literature has also proposed cost-related IVs based on legally protected areas as a proxy for environmental costs (Medeiros *et al.*, 2021b) and demographic variables – for instance, populational density – to represent physical human costs as expropriation and interferences (Zhang e Ji, 2019). We are not arguing that those kinds of costs are equal, but they have similar attributes in how they act as some of the main highway infrastructure costs.

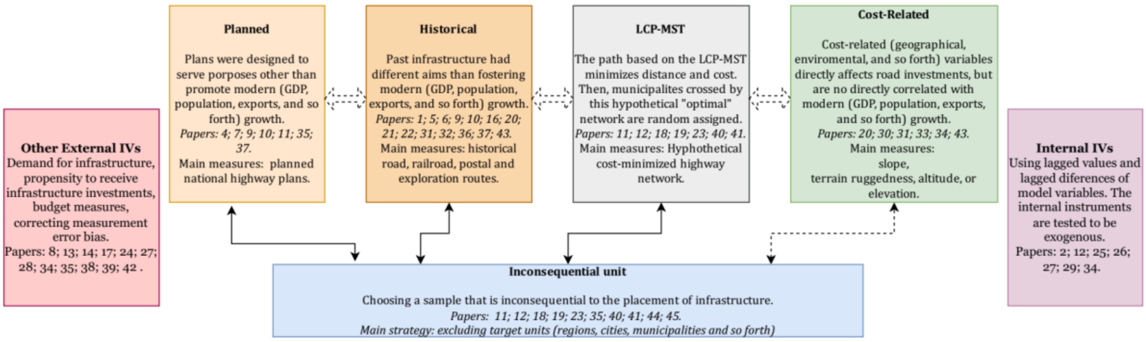
We can also identify another disconnected group of studies using different external IVs with distinct rationales. For instance, Coşar and Demir (2016) used the initial share of expressways as an instrument for provincial access to gateways. Fedderke and Bogetic (2009) constructed a demand for infrastructure IV based on the predicted demand for infrastructure stocks using time series econometric models. Baum-Snow (2007) combined the planned route IV with national construction rates over time as an IV. Storeygard (2016) used oil prices to instrument a road cost variable. Gertler *et al.* (2019) relied on IVs of national and provincial road budgets. As we can see, the IVs used in those studies are pretty specific, unlike the other five approaches described before.

Using panel data models, some studies applied internal instruments (lagged values and lagged differenced values of the model variables) in a GMM estimation setting (Albalate and Fageda, 2016; Bird and Straub, 2020; Fahardi, 2015; Fingleton and Szumilo, 2019; Jiwattanakulpaisarn *et al.*, 2010; Ke and Yan, 2021; Liu *et al.*, 2022; Barzin *et al.*, 2018). Some of those studies combine internal with external IVs (Gertler *et al.*, 2019; Holl, 2012; Medeiros *et al.*, 2021a).

Finally, another strand of literature has used an infrastructure reliance measure as an identification strategy (Li *et al.*, 2017; Medeiros *et al.*, 2021b; Percoco, 2015; Wang, Wu, and Feng, 2020). The rationale behind this strategy is that industries that, for technological reasons, tend to rely less on transportation services (e.g., because they move lighter goods that tend to be less "road intensive" (Duranton *et al.*, 2014) or because their employees do not need to travel long distances for business trips) act as a sort of control group for the "treated" ones, i.e., those that rely more on transportation services. This approach is particularly suitable when using sector or firm-level data. Some investigations have used IVs to correct potential measurement errors in infrastructure reliance measures (Li *et al.*, 2017). Medeiros *et al.* (2021a) combine the infrastructure reliance approach with internal and external IVs to estimate the road impact on productivity in Brazilian economic sectors.

Figure 2.1 summarizes the main identification approaches – focusing on those using IVs – used by the infrastructure-development literature. The arrows indicate some convergence between some identification approaches – for instance, some studies used both historical and cost-related IVs, or planned and historical IVs, or some combination of the inconsequential unit approach with some external IV, and so forth. On the other hand, in most cases, the studies classified as "other external IVs" propose specific IVs and are not much related to different frameworks. Similarly, the internal IVs, combined with external IVs in a few cases, are mostly unrelated to the historical, planned, LCP-MST, or cost-related approaches.

Figure 2.1. Related literature: main IV approaches used in infrastructure-regional economic development studies



1. Adler et al. (2020) | 2. Albalate and Fageda (2016) | 3. Barzin et al. (2018) | 4.Baum-Snow (2007) | 5. Baum-Snow et al. (2017) | 6. Baum-Snow et al. (2020) | 7. Bird and Straub | 8. Coşar and Demir (2016) | 9. Duranton and Turner (2012) | 10. Duranton et al. (2014) | 11. Faber (2014) | 12. Fahardi (2015) | 13. Fedderke e Bogetic (2009) | 14. Fingleton and Szumilo (2019) | 15. Frye (2016) | 16. Garcia-López et al. (2015) | 17. Gertler et al. (2019) | 18. Ghani, Goswami and Kerr (2014) | 19. Herzog (2021) | 20. Holl (2012) | 21. Holl (2016) | 22. Hsu and Zhang, 2012 | 23. Huang and Xiong (2018) | 24. Jaworskiy and Kitchensz (2019) | 25. Jiwattanakulpaisarn et al. (2010) | 26. Ke and Yan (2021) | 27. Lee (2021) | 28. Li et al (2017) | 29. Liu et al. (2022) | 30. Lu et al. (2022) | 31. Martın-Barroso, Nunez-Serrano and Velazquez (2015) | 32. Martincus et al. (2017) | 33. Medeiros et al. (2021a) | 34. Medeiros et al. (2021b) | 35. Michaels (2008) | 36. Percoco (2015) | 37. Rokickia and Stępniak (2018) | 38. Storeygard (2016) | 39. Wang, Wu and Feng (2020) | 40. Xu and Feng (2022) | 41. Yang (2018) | 42. Zhang e Ji (2019) | 43. Zhang, Hu and Lin (2020) | 44. Chandra and Thompson (2000) | 45. Banerjee et al. (2012)

Source: authors' elaboration.

2.2.3. Open points and potential progress

The literature using IV identification strategies has provided important advancements in identifying the causal impacts of highway investments on local economic activity. However, we can raise some gaps in the literature that need careful attention.

First, strong instruments are hard to find in practice. This issue comes from the lack of data in many cases, especially in developing countries, wherein more disaggregated data on highways is poor. In addition, while planned routes, historical networks, and LCP-MST IVs have been proven to be strong predictors of road measures in many applications, their replicability might be limited. In most countries around the world, there is no information on historical routes or past road plans we could rely on to develop a potential exogenous instrument. Regarding LCP-MST IV, it might be pretty tricky to identify cities (hubs) targeted by planners. In developing countries, road projects often target areas with high transport demand, but the absence of strategic government planning hinders the identification of potential hubs and the creation of artificial networks applying the LCP-MST method.

Second, measurement error bias from highway measures is often overlooked. This issue is critical when using monetary measures as investment or expense flows, as long as they likely embed infrastructure projects' inefficiencies, corruption, and long-term mature nature. While several studies have argued against using monetary variables (Calderón and Servén, 2014; Straub, 2011), highway investment flows constitute a direct way to measure infrastructure profitability and provide an easy-to-understand indicator for planners and the society. Also, monetary variables capture quality-related investments, which is absent in several applications using physical measures such as road length or access. Finally, the number of granular datasets and information about infrastructure investment allocation has grown worldwide in the last decades due to the increasing demand for transparency in public administration. Using this data is an obvious way to evaluate the efficiency, effectiveness, and cost-benefit of public spending (Brooks and Liscow, 2019; Kornfeld and Fraumeni, 2022). Then, providing a robust identification strategy dealing with measurement errors in monetary road variables is an essential contribution to the empirical literature.

Third, empirical strategies that deal with different road-building types are scarce. Governments might invest in constructing new roads in poorly connected or isolated regions or improving existing roads in localities with a higher expected return to transportation infrastructure. Those different kinds of interventions have particular costs and likely impact economic activity heterogeneously. It implies that the economic return of road investment is diverse across the space. Therefore, if data is available to conduct this type of analysis, identification strategies handling road intervention heterogeneity are critical to provide novel and reliable results on causal road impacts on the local economy.

Finally, an important and not widely used *ex-ante* step in the identification strategy is to rely on a study design that alleviates endogeneity concerns. Some interesting cases can be found in the literature evaluating the causal impact of national highway policies on local outcomes (Baum-Snow *et al.*, 2017; Bird and Straub, 2020; Faber, 2014; Herzog, 2021). Those studies evaluate the impact of road investments made by high-level governments on economic activity in smaller units such as districts, cities, and municipalities. This design reduces endogeneity issues as we do not expect small localities to influence national government decisions directly. Then, the policy exogeneity is more plausible than studies evaluating road policies impacts on economic activity at the same administrative level. In addition, it allows us to work with a substantially larger number of observations, even when excluding central cities, using the inconsequential unit approach.

In the next section, we describe our novel identification strategy, trying to solve the issues mentioned in detail. In a study design evaluating the impact of a national road program on local economic activity, we combine cost-related instruments fixing measurement error bias in highway investment variables with LCP-MST, political, and propensity to receive road intervention instruments correcting the non-random placement of road bias. Whenever feasible, we construct IVs that we believe are quite replicable worldwide. By doing so, we provide a reliable empirical approach to identifying road impacts on local outcomes. In addition, our strategy allows studies to properly work with monetary highway variables, an issue that has been largely neglected so far.

2.3. Measuring causal impacts of road investments on local outcomes: a new three-step proposal combining cost-related and non-random placement IVs

To address the critical issue of measuring the causal impacts of road investment on economic activity, we develop a novel empirical strategy in a three-step setting. Our main goal is to provide a reliable identification strategy that is robust to measurement error and omitted variable bias.

First, we describe our study design evaluating the impacts of a national-level highway policy on local outcomes. To this end, we rely on a national road program (a program within the PAC) implemented by the Brazilian Federal Government between 2007 and 2018. By desegregating national investments at the municipal level, the procedure we describe in detail in the data section, we can evaluate the local impacts of an extensive and aggregated road policy. This study design is essential to alleviate endogeneity concerns between infrastructure investments and economic activity (Faber, 2014; Herzog, 2021), as we do not expect municipalities to directly influence the Brazilian federal government's decisions to place highways across the country. A more detailed description of the motivation behind this first step can be found in Appendix A.

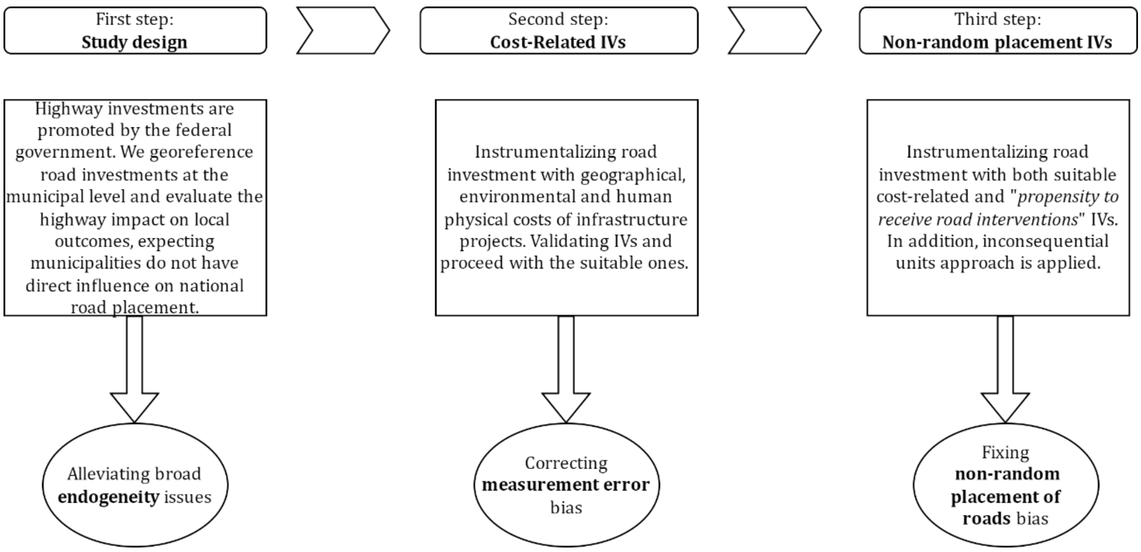
Second, we correct measurement errors from inefficiencies in road projects using the main related geographical, environmental, and human physical costs as IVs for highway investments. In this second step, we statistically validate several replicable cost-related IVs and use them in the succeeding step. This kind of IV may avoid (or alleviate) endogeneity bias in two ways. First, they may solve endogeneity issues related to omitted variables bias commonly found in infrastructure-economic development studies, as proposed by past studies (Holl, 2012; Lu *et al.*, 2022; Martín-Barroso, Nunez-Serrano, and Velazquez, 2015; Zhang, Hu, and Lin, 2020). Second, which we consider more reasonable, they might act as a corrective instrument for measurement error bias of highway investment variables. This is particularly relevant when using monetary highway variables as investment flows in developing economies, wherein a high inefficiency in allocating infrastructure investments is expected (Calderón and Servén, 2014; Straub, 2011). A detailed description of the rationale behind these instruments, as well as Brazilian examples, can be found in Appendix B[4].

Third, we fix omitted variable issues from the non-random placement of highways by combining our previous two steps with IVs related to what we call "the local propensity to receive road interventions." Even when using cost-related IVs to instrumentalize highway investments and potentially correct measurement errors, omitted variable bias from non-random road placement may persist. Correcting measurement errors from infrastructure project inefficiencies is just an (important) part of the problem, but it is likely not enough to eliminate endogeneity bias related to the propensity of certain localities to receive federal highway interventions. First, governments might allocate roads to underdeveloped regions to promote regionally balanced economic growth or direct highway investments to more developed localities – wherein the expected return to road investment is higher – to foster national economic development. Second, political reasons might guide governments to connect the country regionally. Third, by using highway investment flows disaggregated by intervention types – as building, paving, duplications, enhancements and so forth – as we propose in this study, we also need to correct endogeneity bias from the propensity to a locality already connected by a highway in the start period to receive a road intervention. In our third step, we utilize three main IV types to fix those issues. First, we use the LCP-MST global cost minimization process

[4] Figure B1 in Appendix B summarizes the rationale behind cost-related IVs.

(Faber, 2014). To calculate our hypothetical road network, we identify potential hubs (targeted municipalities) based on the start and end points of the highways receiving PAC interventions, which we believe can be easily replicated for several applications. Then, we generate a hypothetical road network minimizing the Euclidean distance between hubs and geographical, environmental, and human physical costs. To the best of our knowledge, no study has applied the LCP-MST minimizing environmental and human physical costs. Second, we also try IVs based on Brazilian historical plans (Bird and Straub, 2020), which alleviate concerns about politically biased road allocation. Third, we use the proximity to which we name "potential road intervention areas" to instrumentalize highway investments in municipalities already connected by federal roads in the start period[5]. To this end, we use traffic intensity data to establish cities that are highly likely to receive road interventions, being an original and highly replicable instrument. Complementary to the proposed IVs, we also use the inconsequential unit approach pioneered by Chandra and Thompson (2000) to exclude likely targeted cities. A full description of the rationality behind the non-random allocation IVs can be found in Appendix C. Figure 2.2 sums up our empirical approach.

Figure 2.32. Three-step empirical approach: identifying causal impacts of highway investment on local outcomes



Source: authors' elaboration.

2.4. Data

2.4.1. National highway investments

To measure the impact of national highway investments on local (municipal) economic activity, we construct a new dataset of national investment flows at the municipal level from 2007 to 2018. To this end, we use two main publicly available

[5] In these cases, buildings are related to lane duplications and road improvements instead of construction and paving.

datasets. The first one concerns data on investment flows in the highway sector of the Federal Government's Growth Acceleration Program (PAC). This dataset includes annual information on road investment flows for each of the 27 Brazilian states, including a brief description of each intervention. The second one refers to the National Highway System (SNV) georeferenced road data made available by the National Highway Infrastructure Department (DNIT). From this data, it is possible to identify the length of each road segment, its condition (paved, duplicated, and so forth), and the places (municipalities) crossed by each one of the intervention road segments. The PAC dataset is not georeferenced, which implies we have restricted information about whether and to what extent an intervention crosses a municipality. To create a national highway investment dataset at the municipal level, we combine the PAC's data description of each intervention and the georeferenced SNV data. The first step was to identify the treated highway codes and their starting and ending points from the intervention description of the PAC data. Next, the PAC-treated highways were linked to the SNV geolocalized data using the highway code and their starting and ending points. Second, we calculated the total PAC intervention road length by municipality and use it to measure the share of the road length in the municipality in relation to the total intervention road length. As we have investment data only by intervention, we use the measured share to compute the highway investment by municipality. It should be noted that the maintenance intervention descriptions barely describe the state and the highway code. In this sense, it was not possible to geolocate this type of investment at the municipal level, and they were excluded from further analysis. Then, our main interest variable is the road investment (R$) at the municipal level. As robustness checks, we will also try two additional road variables. The first one is a dummy variable assuming value one if the municipality received a road investment during the PAC period and zero otherwise. The second one is the road length growth rate between 2006 and 2018. In this case, we use 2006 data from the 2007 National Transport Logistics Plan (PNLT) and 2018 data from DNIT[6]. A detailed description of our infrastructure variables can be found in Appendix D[7].

2.4.2. Infrastructure project cost-related IVs

To construct the cost-related IVS, we use a set of variables representing environmental, geographical, and expropriation costs at the municipal level. At the environmental scope, we use georeferenced data of legally protected areas[8] in the National Registry of Conservation Units (CNUC), maintained by the Ministry of the Environment (MMA). Then, we merge this data with the municipality boundaries shapefile to identify whether a legally protected area intersects a municipality. Our variable is a dummy, which

[6] We can likely observe measurement error in the road length variable as well, as the PNLT and DNIT files are not fully comparable. In addition, there is methodological variations over the years in relation to road classifications as federal, state level and so forth. Then, this variable should be used with caution.
[7] Figures D1, D2 and D3 in Appendix D exhibits our constructed highway investment data at the municipal level.
[8] The data are divided into six groups: Federal Full Protection and Sustainable Use Conservation Units, State Full Protection and Sustainable Use Conservation Units, and Municipal Full Protection and Sustainable Use Conservation Units.

assumes a value of one if a legally protected area intersects the municipality and 0 otherwise. Second, we utilize the environmental embargo terms data of the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) inspection system. Environmental embargoes represent penalties applied to prevent an exploratory activity from continuing. The embargos also serve to inhibit ongoing damage and promote environmental recovery. To generate our variable, we first aggregate the number of embargoes in the five previous years (2002-2006)[9] from the PAC by municipality. Hence, we create a dummy variable that assumes a value of 1 if there was an environmental embargo in the municipality during this period and 0 otherwise.

To construct geographic-related IVs, we rely on a few studies measuring infrastructure effects on local outcomes using this kind of identification approach (Holl, 2012; Lu *et al.*, 2022; Martin-Barroso, Nunez-Serrano, and Velazquez, 2015; Medeiros *et al.*, 2022; Zhang, Hu, and Lin, 2020). All those works utilize some measure based on slope, elevation, ruggedness, or altitude. Our preferred measure is the share of the municipality area with a slope above 20%, which corresponds to hilly regions. This variable is highly related to road construction in the world and Brazil, as DNIT defines maximum values for slopes to be applied to the construction of highways and local roads. The slope is characterized by the relation between a gradient and a corresponding distance on a tiny scale, which is unlikely to affect any development outcome directly at the aggregated municipal level. To calculate this variable, we use slope raster data from the National Institute for Space Research (INPE), which allows us to count the number of slope pixels above the 20% cutoff. Then, we generate the share of hilly pixels in relation to the total pixels as our main geographic IV.

We use urban infrastructure building variables to quantify expropriation (human) costs. Our preferred variable is the share of urban infrastructure[10] building in relation to the total municipality area. For this, we use land use and land cover data made available by MAPBIOMAS (Souza *et al.*, 2020), extracted at the municipal level.

We try several other geographical, environmental, and human costs as robustness checks. We also create composite cost indexes (Cost Index 1 and Cost Index 2) based on dimensionality reduction methods[11], as cost types may have some complementary characteristics. We generate our composite indexes using our preferred measures – legally protected areas, environmental embargos, sloped areas, and urban infrastructure. Cost Index 1 can be interpreted as an environmental cost index, while Cost Index 2 can be

[9] We take the 5-years sum to avoid potential outliers at using annual data.

[10] This variable is related to construction and infrastructure and is used to generated urban density areas measures.

[11] We use Multiple Correspondence Analysis (MCA) for mixed data to generate our composite indexes. We use the first two components as they accumulate 59% of the original data variation. We name the first component as Cost Index 1, and it is calculated by $Cost\ Index\ 1 = 0.58 * Legal\ Protected\ Area + 0.51 * Embargos + 0.13 * Sloped\ Area + 0.08 * Urban\ Infrastructure$. The second component, which we name Cost Index 2, is calculated by $Cost\ Index\ 2 = 0.00 * Legal\ Protected\ Area + 0.01 * Embargos + 0.49 * Sloped\ Area + 0.58 * Urban\ Infrastructure$. We expect the Cost Index 1 to be more plausibly exogenous, as it received lower influence from the urban infrastructure variable.

seen as a geographical-expropriation index. We describe our preferred and robustness checks variables and their sources in detail in Appendix E.

2.4.3. Non-random road allocation IVs

We use three sets of instruments to represent the "local *propensity to receive road interventions."* The first ones are constructed using the LCP-MST method. To create our LCP-MST hypothetical road system, we first need to point out the hubs that the network should connect through a minimization cost process (Dijkstra, 1959; Kruskal, 1956)[12]. Our preferred strategy considers the starting and ending points of highways receiving PAC investments as hubs. To identify those central cities, we use SNV georeferenced data and generate a dummy variable assuming value one if a municipality is a starting or an ending point of a PAC intervention road and zero otherwise. This approach gains relevance, especially in country contexts where road policies have no clear direction, making it hard to predict which places governments aim to connect. We also try to establish hubs based on the centrality of cities and historical population data. In addition, we calculate the LCP-MST network, minimizing both Euclidean distance and an infrastructure cost index based on our preferred geographical, environmental, and human physical variables, allowing us to capture better likely paths for road infrastructure interventions aimed at reducing overall costs. We describe this procedure and the following non-random allocation IVs in detail in Appendix F. Our preferred LCP-MST network can be seen in Figure F2 in Appendix F.

The second ones are based on the Brasília Plan. We rely on Bird and Straub (2020) to construct political IVs based on historical plans. Bird and Straub constructed a hypothetical radial (straight line) network connecting the capital, Brasília, to eight critical cities around the country. By linking the capital to those cities, the radial network established corridors, which incidentally connected other localities along the way. Our first political IV is measured as the distance from a municipality center to the nearest Brasília Plan segment. We replicate the index proposed by Bird and Straub (2020) based on buffer zones around the straight lines and the shares of each municipality's area within each zone. Our second political IV is based on the Juscelino Kubitschek (JK) Road Cruise, an extension of the Brasília Plan. We digitalized old maps to construct our second historical instrument. We follow the procedure in Bird and Straub (2020) and apply it to the JK Cruise network, generating 60 km buffer zones around the lines (see Figure F5 in Appendix F). Then, our two political IVs are distance-based indices from the municipalities' centers to the lines weighted by the municipal area shares in the buffer zones.

The third one is based on traffic intensity data. First, we use 2007 PNLT data on traffic intensity to identify "*potential road intervention areas*." In this data, road segments

---

[12] Following Faber (2014), we use the Dijkstra's (1959) optimal route algorithm to compute least costly construction paths between any bilateral pair of targeted nodes. Then, we use these bilateral cost parameters in combination with Kruskal's (1956) minimum spanning tree algorithm to identify the subset of routes that connect all targeted nodes on a single continuous graph subject to global network construction cost minimization.

are based on the classification A-F of traffic intensity commonly used in transportation engineering studies in Brazil. Based on the vehicle flows on federal roads, highway segments are classified from A (light traffic) to F (heavy traffic, including huge traffic jams). Then, we consider those segments classified as D, E and F as heavy traffic roads, being of potential federal government focus. The heavy traffic areas can be visualized in Figure F6 in Appendix F. Next, our IV is calculated as the distance from the municipality center to the nearest heavy-traffic road segment. The rationale behind this IV is that, conditional on controls, municipalities already connected by federal roads in 2006 and located close but not so close to "*potential road intervention areas*" are more likely to (inconsequentially) receive highway investments to reduce traffic levels in the critical areas. In this case, including demographic controls is an essential condition to ensure IV validity, as we can expect heavy traffic roads crossing urban agglomerations to be directly impacted by events involving people, such as accidents paralyzing roads, the need for traffic signals reducing car travel speed, and so forth.

2.4.4. Dependent and control variables

Our primary dependent variable is the Gross Domestic Product (GDP) *per capita*. We chose this based on the related literature, in which we observed a massive use of GDP as one of the interest variables. In addition, GDP per capita can be used as a proxy for productivity, which allows us to calculate the return rate to highway investments and provides a more interpretable result for policy purposes. As robustness checks, we also use employment, firms, and wages as labor market measures.

To avoid omitted variable bias, we include an extensive set of controls (see Appendix G for more details about the motivation to include those variables and their sources and formulas). First, we include the initial level of the dependent variable as a control for the municipality development level. We also include the share of poor people to control for policies oriented to poverty alleviation, which is a characteristic of most of the PAC period. Second, as road infrastructure is served by the federal and state level governments, we include state-fixed effects to control regional infrastructure policies. Third, the municipality area is included to control for territorial size. This control is critical as our main infrastructure variable is based on the PAC road length crossing a municipality area. As an additional municipal size variable, we include the formal workforce measured as the log of the number of formal workers in 2007. Fourth, we include one important control related to the agricultural sector. Fifth, we include the share of exports of each municipality in the national exports in 2007 as a control. Sixth, we include a set of controls related to complementary infrastructures. Then, we include the distance (km) to the nearest port, railroad, and state road in 2006 as controls. Seventh, we control for the historical propensity of a municipality to receive federal road investments. Our variable is the number of railway stations in 1920, the main transportation sector in that period. Finally, we include some controls related to the municipal social and institutional background. We first include the Index of Municipal Institutional Quality (IQIM) to control for municipal institutions. Second, we include the population share with master's or doctoral degree as a proxy for local social development. Table G1 in Appendix G summarizes our variables and their respective data sources. Table G2 shows descriptive statistics.

2.5. Econometric specification

In this section, we present our empirical approach. The *ex-ante* first step consists of our study design based on disaggregated highway investment data, which evaluates the impact of a national road program on local (municipal) outcomes and alleviates broad endogeneity issues between road investment and economic activity. From that, we apply the next two steps, which we describe in detail below.

2.5.1. Correcting measurement error bias (the second step): testing cost-related IVs

Our second step uses the main geographical, environmental, and human physical costs of infrastructure projects as instruments for our highway investment measure. As we described before, the rationality of this step is to correct measurement error bias from several inefficiencies in road investments, such as poor project design. Those inefficiencies lead to measurement errors, as highway investment flows do not fully represent the available road infrastructure for use by the population.

To overcome this issue, we test for suitable cost-related IVs for highway investments. Our main objective is to estimate the long-term effect of improvements in highway network on our outcome variables at the municipal level. Our main second stage equation is given by:

$$\Delta Y_{is} = \beta_0 + \beta_1 \Delta HighwayInvestments_{is} + \beta_2 X_{is} + \theta_s + \varepsilon_{is}, \tag{1}$$

Where $HighwayInvestments_{is}$ is the log form of the federal highway investments, β_0 is a constant term, β_1 is the elasticity of highway investment to our dependent variable, Y_{is} is our dependent variable for municipality *i* in state *s*, X_{is} is a vector of control variables, θ_s is a state fixed effect, β_2 is a vector of parameters related to the control variables and ε_{is} is the idiosyncratic error term.

The term Δ represents the change in our dependent variable between 2007 and 2018, i.e., ΔY_{is} is the GDP *per capita* growth rate. As our highway variable measures investment flows, the highway change is the sum of the investment flows between 2007 and 2018.

In our second step, the corresponding first-stage equation is given by:

$$\Delta HighwayInvestments_{is} = \beta_3 + \beta_4 CostRelatedIV_{is} + \beta_5 X_{is} + \theta_s + \epsilon_{is}, \tag{2}$$

where $CostRelatedIV_{is}$ is a vector containing our cost-related IVs, and β_3, β_4 and β_5 are parameters to be estimated. We run equations 1 and 2 using two-stage least squares (2SLS) and test for IV strength using robust first-stage F-statistics (Imbens, 2014). This approach allows us to identify suitable cost-related IVs for highway investments and, hopefully, correct measurement error bias resulting from infrastructure project inefficiencies.

2.5.2. Fixing omitted variable bias (the third step): including non-random allocation IVs

In our third step, we include a set of "*local propensity to receive road interventions*" (non-random road placement) IVs. Then, we rewrite the first-stage equation as follows:

$$\Delta HighwayInvestments_{is} = \beta_6 + \beta_7 CostRelatedIV_{is}^{suitable} + \beta_8 NonRandomAllocationIV_{is} + \beta_9 X_{is} + \theta_s + \epsilon_{is}, \tag{3}$$

where $CostRelatedIV_{is}^{suitable}$ is a vector of suitable cost-related IVs tested in the second step, $NonRandomAllocationIV_{is}$ is a vector of non-random allocation IVs, and β_7 and β_8 are they respective parameter vectors.

As we described earlier, the vector of non-random allocation IVs includes several LCP-MST, political, and "*potential road intervention areas*" measures. The first two kinds are more related to constructing new roads, as they predict a hypothetical road network based on global cost minimization or political reasons. On the other hand, our "*potential road intervention areas"* IV works for municipalities crossed by a federal road in the start period, i.e., they are strictly related to duplications or enhancements of existing roads. To capture those heterogeneities in road investment, we expand our third step first-stage equation as follows:

$$\begin{aligned} \Delta HighwayInvestments_{is} = \beta_{10} + \beta_{11} CostRelatedIV_{is}^{suitable} + \\ \beta_{12} D_{is}^{New} \times NonRandomAllocationIV_{is}^{LCP-MST} + \\ \beta_{13} D_{is}^{New} \times NonRandomAllocationIV_{is}^{Political} + \\ \beta_{14} D_{is}^{Existing} \times NonRandomAllocationIV_{is}^{InterventionAreas} + \\ \beta_{15} X_{is} + \theta_s + \epsilon_{is} \end{aligned} \tag{4}$$

where D_{is}^{New} is a dummy variable assuming value 1 whether a municipality were not connected by a federal road in 2006, and zero otherwise, $D_{is}^{Existing}$ is a dummy variable assuming value 1 whether a municipality were crossed by a federal road in 2006, and zero otherwise, $NonRandomAllocationIV_{is}^{LCP-MST}$ is a vector of LCP-MST IVs, $NonRandomAllocationIV_{is}^{Political}$ is a vector of political based IVs, $NonRandomAllocationIV_{is}^{InterventionAreas}$ is our *"potential road intervention areas"* IVs, and β_{11} to β_{14} are their respective parameters to be estimated.

By running equations 4 and 1, we expect to correct both measurement error and omitted variables bias of road investments. In addition, by differentiating for new and existing roads in the start period, we provide a novel strategy to estimate the causal impact of highway investments in a multitype road intervention setting.

2.6. Empirical results

2.6.1. Main results

We start our econometric analysis by running the second-step models. Appendix H contains a complete description of the second-step results, correlation matrices, and OLS estimations used for comparison (see Figure H1 and Tables H1-I4 in Appendix H). In Table H5 in Appendix H, we propose a set of specifications based on our preferred cost-related IVs. To check instrument strength, we report KP Wald F and *effective* F (Olea and Pflueger, 2013) statistics. Both statistics are robust to heteroskedasticity and weak instruments, and the *effective* F statistic works suitably even in multiple instrument settings (Andrews, Stock, and Sun, 2019), as proposed by our identification strategy. The first stage regressions show that our preferred cost-related IVs strongly predict the long-term changes in the national highway investments at the municipal level. Nonetheless, some other cost-related IVs seem to violate exclusion restriction or unconfoundedness. Then, we proceed to the third step with five suitable instruments: the legally protected area, environmental embargoes, sloped area, Cost Index 1, and Cost Index 2.

By applying our second-step identification strategy, we fixed (or alleviated) the expected measurement error underestimation bias. In this sense, our second step parameters may be understood as "*free from measurement error*" elasticity. Nonetheless, two obvious empirical issues remain. First, some measurement error bias might remain if our instruments do not fully or genuinely capture the main infrastructure costs, leading to inefficiencies. This issue seems more troublesome for human physical IVs, which might correlate with socioeconomic variables affecting productivity and infrastructure placement. Second, non-random allocation bias might exist even after correcting for measurement errors in highway variables.

Now, we evaluate whether the highway investment impacts found so far remain unchanged when correcting for both measurement error and non-random road allocation bias using our proposed LCP-MST, historical, and intervention area IVs. We use our preferred second-step (and hopefully "free from measurement error") specifications in columns 10 and 11 of Table H5 in Appendix H as our starting point for the third-step regressions. Then, we test several specifications using our non-random allocation IVs.

Table 2.1 summarizes our third step and final econometric results. We start by inserting our preferred[13] LCP-MST IV using starting and ending road points as hubs (columns 1 and 2). Next, we try our "*potential road intervention areas*" IV (columns 3 and 4) and Brasília Plan (columns 5 and 6) separately. In columns 7 and 8 we include our full set of non-random allocation and suitable cost-related IVs as proposed in equation 4. In columns 9 and 10, we test a Non-random Allocation Index composed of our preferred LCP-MST, historical, and intervention areas IVs[14]. Finally, we use the inconsequential unit

[13] We also try LCP-MST IVs based on REGIC and historical cities as robustness checks (Table I1 in Appendix I). Results remain stable. We also try the JK Road Cruise IV as a robustness check for historical IV.

[14] To construct our index, we use Principal Component Analysis (PCA). The Non-Random Allocation Index is measured as follow: $Nonrandom\ Allocation\ Index = 0.58 * LCP|MST + 0.58 * BrasíliaPlan - 0.57 * PotentialInterventionAreas$. We use the first principal component, presenting 89% of cumulative data variation.

approach in all specifications, excluding hubs, historical cities, and/or critical traffic points, which would likely cause selection biases.

All our proposed non-random road allocation IVs strongly correlate with the highway investment variable. First-stage results show a consistent correlation between instruments and highway investments and suggest that our identification strategy seems solid at identifying the causal impacts of road investments on local economic activity. *Effective-F* statistics severely increase by including non-random allocation IVs compared to second-step estimates, even after solving (or alleviating) measurement error bias. In this sense, our findings restate the critical role in fixing endogeneity bias from the non-random placement of highway investments, especially using monetary variables. In addition, results indicate that our full econometric strategy following equations 1 and 4 is suitable for measuring causal road impacts on local productivity in the context of road intervention heterogeneity. In contrast, findings put some caution at using only cost-related IVs in contexts of non-random road policy allocation and seem to indicate cost-related IVs solving (or alleviating) just a (relevant) part of the problem.

Now, we turn our analysis to the IV parameter's direction and magnitude. Starting with our LCP-MST IV, we find an expected negative relationship with highway investment, indicating that municipalities far from the hypothetical LCP-MST network received a smaller amount of road investments. The same holds for our Brasília Plan IV, and the interpretation is quite similar, suggesting that localities far from the plan lines received lesser investments. On the "*potential intervention road areas*" IV, the positive parameters suggest that, for already highway-connected municipalities in 2006, the greater the distance from a critical traffic area, the greater the investment received. This result points out that planners seem to avoid critical road segments by directing interventions to road segments not so close to critical areas. It might be related to the potential correlation between population density, other human physical costs, and critical traffic points, which is likely impacted by higher economic costs and inefficiency in those segments. Then, critical road segments appear to be a strategic (and likely endogenous) feature in road policies and using the inconsequential unit approach is overarching in this way.

Regarding second-stage results, elasticities are stable from 0.011 to 0.017. In other words, a one-percent increase in federal highway investment boosts municipal GDP per capita by 0.011-0.017 percent. It is important to note that the third-step elasticities are not negligibly lower than those in the second step (between 0.02 and 0.03). This result suggests that even correcting for measurement error bias, a non-random allocation overestimation bias might remain, and our identification strategy is essential to rule those issues out.

Table 2.1. Federal Highway Investments and Municipal GDP *per capita* Growth, 2007-2018: Third Step 2SLS IV Regressions

	1	2	3	4	5	6	7	8	9	10
Second stage										
Log Highways Investments	0.0165***	0.0162***	0.0126**	0.0122***	0.0110**	0.0110**	0.0124**	0.0124***	0.0129***	0.0127***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
First stage										
Legal Protected Areas	0.4597***		0.3555***		0.4232***		0.3389***		0.3423***	
	(0.10)		(0.10)		(0.10)		(0.10)		(0.10)	
Environmental Embargos	0.3312***		0.3478***		0.3547***		0.3383***		0.3384***	
	(0.11)		(0.11)		(0.11)		(0.11)		(0.11)	
Sloped Area	-1.3606***		-1.6060***		-1.4378***		-1.4601***		-1.4676***	
	(0.31)		(0.31)		(0.32)		(0.30)		(0.29)	
Cost Index 1		0.2586***		0.2274***		0.2389***		0.2177***		0.2190***
		(0.05)		(0.05)		(0.05)		(0.05)		(0.05)
Cost Index 2		0.3571***		0.4427***		0.3756***		0.3882***		0.3896***
		(0.07)		(0.07)		(0.07)		(0.07)		(0.07)
LCP-MST Starting and Ending Road Points	-0.3875***	-0.3918***					-0.1193***	-0.1204***		
	(0.02)	(0.02)					(0.04)	(0.04)		
Potential Road Intervention Area			0.3252***	0.3294***			0.0893*	0.1025**		
			(0.02)	(0.02)			(0.05)	(0.05)		
Brasília Plan					-0.3170***	-0.3176***	-0.1421***	-0.1314***		
					(0.01)	(0.01)	(0.03)	(0.03)		
Non-Random Allocation Index									-0.4994***	-0.5035***
									(0.02)	(0.02)
Observations	5402	5391	5190	5178	5469	5457	5126	5115	5126	5115
KP Wald F Statistic	113.896	156.650	97.221	137.336	123.195	169.934	72.567	89.717	106.925	147.383
Effective F Statistic	104.053	117.056	96.535	108.326	108.318	122.313	73.324	78.795	104.841	112.493
2SLS critical value for tau=5%	20.184	21.202	19.398	20.783	20.688	21.922	24.187	26.112	20.345	21.008
R^2	0.22	0.22	0.24	0.24	0.23	0.23	0.23	0.23	0.23	0.23

All regressions include the following set of control variables: GDP per capita in 2007; state fixed effects; municipality area; workforce; agriculture share; exports share; distance to the nearest state road; distance to the nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors are reported in parentheses. * 0.1 ** 0.05 *** 0.01.

On the direction of the bias, an analysis through the PAC is quite difficult. The program had several goals, including connecting underdeveloped regions to promote regionally balanced growth and attending developed regions with restrained demand for transport investments to foster national economic development. Complementary, those motivations might be mixed up with political, pre-determined historical infrastructure, and touristic reasons, which is rather hard to distinguish.

Nonetheless, we can provide some elucidation starting from our second-step estimates (Table H5 in Appendix H). Suppose that our preferred second-step regressions entirely correct for measurement error bias but do not so for non-random allocation bias. Then, we have a (potentially biased) start point elasticity around 0.02 and 0.03, which we compare with our smaller third-step elasticities (0.011-0.017). The apparently overestimated second step parameters suggest that, "*free from measurement error bias*", PAC planners targeted comparatively more places with higher expected returns to infrastructure investments and higher expected traffic demand (Baum-Snow *et al.*, 2017; Faber 2014). It seems accurate, as almost half of the PAC highway investments were allocated to road enhancements and duplications in already connected municipalities in 2006. Regarding the other half, we expect a mix of underdeveloped-oriented policies – especially in the poorly connected North region – and higher expected returns on infrastructure policy, mainly in export and agriculture-specialized regions. Medeiros *et al.* (2021a) found that highway investments have higher returns for low-technology-intensive sectors such as agriculture and mining in Brazil. In this sense, the PAC's economically more developed-oriented nature seems stronger than its regionally balanced growth aims, upward biasing estimates "*free from measurement error bias*" but not solving non-random allocation issues. Indeed, this interpretation is more suggestive than a statement, as we do not entirely know whether cost-related IVs are solving some share of the non-random allocation bias or if they are entirely fixing measurement error bias. We continue this discussion in section 6.

2.6.2. Falsification test and robustness checks

In this section, we present several robustness checks to our main results shown in the previous topic. First, we propose a falsification test based on a novel planned road sample exercise. Second, we try other commonly used dependent variables. Third, we test two different measures of road infrastructure. Fourth, we run an additional robustness check using the limited information maximum likelihood (LIML) estimator, bootstrap standard errors, and excluding potential outliers.

If our instruments are valid, they should affect the outcome only through the highway investment variable. Therefore, cost-related and non-random road allocation IVs should not affect local economic outcomes not in the highway investment pathway. To test the exclusion restriction, we specify a sample closely related to our study population but not receiving highway investments. This sample is characterized by municipalities presenting planned roads, i.e., hypothetical highways acting as a guideline intended to meet a potential traffic demand. In this sense, our rationality behind this falsification test is that places (hypothetically) crossed by planned roads were similarly demanding road infrastructure interventions, likely exposed to the same potential confounders as the full sample. As none of those municipalities received federal highway interventions, the falsification test is performed by including the IVs in an alternative specification of the outcome equation. Appendix J describes in detail the results of the falsification tests and all the subsequent robustness checks. We run the same specifications as in Table 2.1 using the planned roads sample and include our preferred cost-related IVs separately (columns 1 and 2). Results can be seen in Table J1 in Appendix J. Conditional on controls, results suggest that our suitable cost-related and non-random allocation IVs do not violate exclusion restriction. These results increase confidence in our identification strategy.

Second, we estimate the impact of federal highway investments on employment, firms, and wages. Results are summarized in Table J2 in Appendix J. Cost-related and non-random allocation IVs strongly predict national highway investments for all tested dependent variables. In addition, the signal and significance of the first-stage coefficients remained relatively similar, and the road impact on outcomes remained positive.

Third, we try two additional infrastructure variables as robustness checks. Several studies correctly argue that monetary variables - such as the investment flows used in our study - may contain several measurement errors (Calderón and Servén, 2014; Kenny, 2009; Straub, 2011). Our dummy variable assumes value one if a PAC highway intervention crossed the municipality and zero otherwise. If there is too high measurement error in our preferred investment flow variable, the intervention dummy variable might alleviate the problem as it no longer contains monetary values. Second, we try a road length variable following a vast strand of literature. Results are described in Tables J3 and J4 in Appendix J. As expected, the highway intervention parameters are positive and significant in all specifications, and the same holds for road length. The cost-related and non-random allocation IVs work the same way as using continuous highway investment flows, corroborating previous estimates.

Fourth, we try additional robustness checks to raise confidence in our main estimates. We run the same specifications of Table 1 using the LIML estimator (Anderson and Rubin, 1949). In addition, we also provide bootstrap confidence intervals. Young (2022) finds that bootstrapped confidence intervals perform better in real-world settings as heteroscedasticity and weak IV assumptions are likely violated. Tables J5 and J6 exhibit the results. The findings remain unchanged in both cases, indicating that our main estimations are reliable.

Finally, we try several specifications, excluding potential outliers. Results are described in Table J7 in Appendix J. In column 1, we drop all municipalities of the state of São Paulo. São Paulo is the most prosperous Brazilian state, representing over 30% of the national GDP. In addition, São Paulo presents the country's best road infrastructure, which is substantially privately managed. A significant share of those high-quality roads is the responsibility of the São Paulo state government, with federal roads a small fraction of the total. The result might be a small fraction of the PAC highway investment directed to a high road-demanding state, and an underestimation bias could be expected. In column 2, we exclude all municipalities in the states in the North region. Those municipalities are characterized by large territorial areas, which might bias our highway measure variable as it depends on the road length crossing the municipalities. In column 3, we exclude both São Paulo and Northern municipalities. In column 4, we consider highway investment values smaller than R$ 50 million to be zero. This test is essential as we relied on road length crossing municipal areas to construct our highway flow measure, and short road segments (and consequentially small investment values) might be poorly capturing a highway intervention. Finally, in column 5, we exclude municipalities in the top 1 and bottom one percentiles of GDP *per capita* growth. In general, findings remain almost unchanged. The most noticeable variation comes from the exclusion of São Paulo municipalities. The elasticity in columns 1 and 3 is around 0.017, while our benchmark estimate is 0.012. This result suggests that São Paulo municipalities might be slightly downward biasing our estimates.

2.7. Assessing the return rate to highway investments: how large can the bias be?

Now, we return to our main issue: endogeneity bias. We can ask some questions. First, is there any remaining bias in our second-step elasticity? If there is, what is the magnitude of such a bias? To answer these questions, we compare our preferred third-step estimates with second-step and OLS naïve estimates (Table H5 and Table H6 in Appendix H, respectively).

The OLS estimates returned an average highway investment elasticity of around 0.004, while the average second-step elasticity is close to 0.025. On the other hand, the highway investment elasticity estimated using our full third-step IV identification strategy is between 0.011 and 0.017. Our results suggest a meaningful bias in OLS regressions and point to a remaining non-random allocation bias in our second-step estimates.

To better illustrate the bias, we calculate the return rate to highway investments in Brazil comparing the OLS and our preferred 2SLS-IV estimates. To do so, we follow the equation below (Fernald, 1999; Wang, Wu, and Feng, 2020):

$$RR = \beta_{highway} * \frac{GDP}{Highway\ Stock} \tag{5}$$

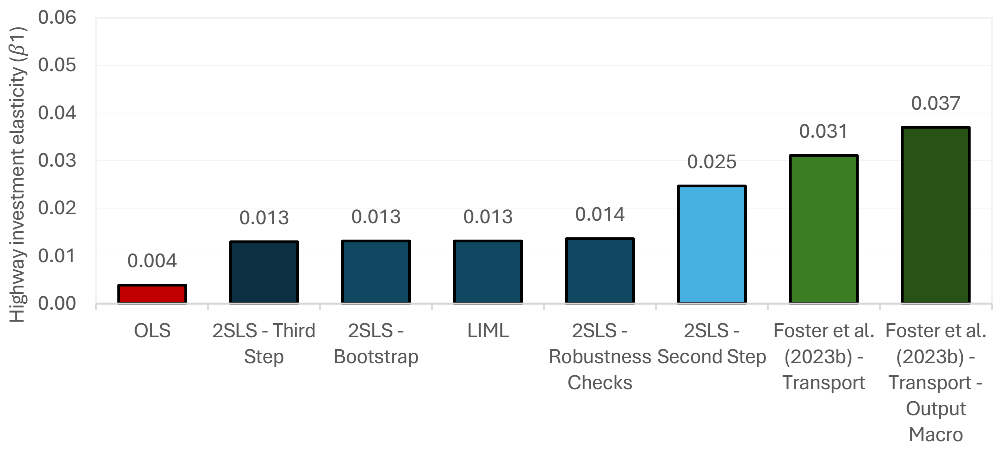
Where RR is the return rate to highway investments, $\beta_{highway}$ is the predicted elasticity following equations 1 to 4 in the econometric specification section, GDP is the national GDP, and Highway Stock measures, in monetary terms, how much the national highway infrastructure stock is worth. As a proxy for the highway stock, we follow Medeiros *et al.* (2021a) and use the sectoral infrastructure stock estimated by Frischtak and Mourão (2017). The study provides a stable road stock of 6.0% of the GDP. The Brazilian GDP was around R$ 9.9 trillion in 2022. Considering the road stock share in GDP constant over time, the actual Brazilian highway stock is supposed to be approximately R$ 594 billion, implying a GDP/Highway Stock ratio of roughly 16.7. This ratio is very close to the Chinese one of 19.6, used by Wang, Wu, and Feng (2020).

In Figure 2.3, we summarize our results in terms of elasticities and return rates. Then, Figure 2.3 shows us the elasticity and return rate by estimation group, which is measured by the average considering all their respective estimates. In addition, we include Foster *et al.* (2023b) transport sector elasticities as benchmarks for comparison as well as the return rates found by Li *et al.* (2017) and Wang, Wu, and Feng (2020) for China, and Medeiros *et al.* (2021) for Brazil.

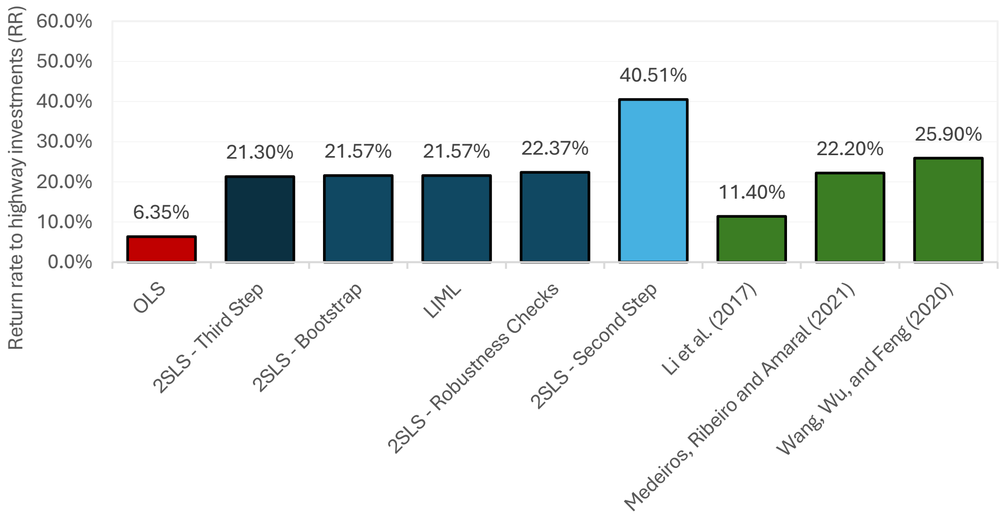
Using the naïve OLS highway elasticity, we generate an underestimated RR of 6.35%. In contrast, using the 2SLS estimates based on the second step cost-related IVs identification strategy, we calculate an RR of around 40.51%, which we consider relatively high even for a developing and lacking highway infrastructure economy such as Brazil. Nonetheless, when we turn to our third-step estimations – including all robustness checks we have run –we find a stable average elasticity around 0.013, implying a return rate to highway investment of approximately 21.3%. This result suggests that correcting measurement errors and non-random allocation bias is critical to identifying robust infrastructure investment elasticities and return rates. In addition, findings put a lot of caution on instrumentalizing highway investment variables only by cost-related measures in cases wherein road policies are likely guided by economic, political, or need for transportation reasons, as seems to occur in Brazil during the PAC.

While the methodology and data are different and making a full comparison among studies is impossible, our calculated RR is in line with those ones calculated by Li *et al.* (2017) and Wang, Wu, and Feng (2020) for the Chinese case, and especially with Medeiros *et al.* (2021a) for the Brazilian case. To evaluate the profitability of highway investments in Brazil, we can take the Social Discount Rate (TSD) of 8.5% calculated by the Ministry of Economy (2021), broadly used to evaluate infrastructure projects in the country. To reduce our third step rate of return of 21.3% to the threshold of 8.5%, Brazil would need 2.51 times more highways, which implies a road stock of 15.1% of national GDP. This finding corroborates the estimates by Frischtak and Mourão (2017) and Medeiros *et al.* (2021a), predicting an ideal road stock of 13.5% and 16% of GDP, respectively. Considering even lower long-run real rates of return from 4% to 5% worldwide, Brazil would need to invest 4.26 times more in highway infrastructure, strengthening the high profitability of transport investments in the country during the PAC period.

Figure 2.3 Elasticities (a) and return rates to highway investments (b): Brazil, 2007-2018



(a)



(b)

Source: authors' elaboration.

2.8. Concluding remarks

Using a novel three-step IV identification strategy, we evaluated the causal impact of national highway investments on local economic activity in Brazil between 2007 and 2018. We relied on several original infrastructure project costs and non-random road placement instruments as sources of quasi-random variation of observed highway infrastructure to correct for both measurement error and omitted variable biases. We proved that our preferred IVs are conditionally exogenous and strong predictors of road investments. In addition, we argue that a substantial part of our proposed instruments might be replicated in several applications worldwide.

Our main results showed that our empirical strategy is suitable for identifying causal road impacts on the local economy. First, the second step estimates allowed us to correct (or alleviate) measurement errors and determine the expected downward-biased OLS elasticities. Second, third step regressions are critical to fix non-random placement of roads bias even after fixing measurement error. In this case, our findings suggest that the PAC prioritizes economic development over regional balance, as its actions favor more prosperous areas, thus upward biasing "free from measurement error" second step estimates. Results remain unchanged under several robustness checks.

We found a pretty stable road elasticity ranging from 0.011 to 0.017. In other words, a one percent increase in highway investments increased municipal GDP *per capita* by 0.011-0.017%. Using a more easy-to-interpret measure, results point to a return rate to highway investment of around 21.3% in Brazil, which aligns with studies worldwide. Brazil would need to invest 2.5 times more in road infrastructure from this return rate, reaching a proper road stock of 15% of the national GDP. This result proved the high rentability of highway investment in the developing country context.

While we have contributed to the empirical literature on infrastructure and economic development in several ways, some gaps remain. First, it is hard to conclude if our second-step estimates are fully corrected for measurement errors in the road variable. In addition, cost-related IVs could be fixing some part of the expected omitted variable bias, and additional validations might be useful. Proposing other cost-related instruments and statistically testing them is an obvious way to confirm our identification strategy and results. Second, a similar empirical approach could be adapted for panel data models. In this case, we could include other cost types, such as national demand and institutional risks, which are issues we overlooked in the present research as we worked with cross-sectional data. We hope to provide some new evidence on those concerns in further research.

3. **INFRASTRUCTURE, GROWTH, AND REGIONAL DISPARITIES:** looking at the efficiency, redistribution, and equity goals in road investments

Abstract

This paper evaluates the role of transportation infrastructure policies in raising productivity by considering heterogeneities in terms of efficiency, road specialization, redistribution, and equity. Using a third-step instrumental variable identification approach combined with the infrastructure reliance framework, we estimate the causal impacts of road investments on GDP *per capita* in Brazilian municipalities during the Growth Acceleration Program (PAC) period (2007-2018). We find positive impacts of road investments on productivity, which can be translated into a return rate of over 20%. This effect is larger for less-developed, poorly infrastructure-endowed, and more road-specialized municipalities. From this, we provide regional focal points constituting *win-win* (efficient-specialized and redistributive-equative) combinations for road policies. In addition, findings suggest that Brazil could have achieved higher profitability for road investments by targeting redistributive-equative and efficient-specialized interventions. However, policymakers need to be cautious as the impact of road investments on productivity is constrained in highly efficient localities, likely due to institutional and sectoral inefficiencies.

Keywords: highway infrastructure; regional development; redistribution; equity; infrastructure heterogeneity.

3.1. Introduction

A wide range of studies has evaluated the impact of infrastructure investments on economic growth, productivity, trade, employment, decentralization, poverty alleviation and reducing inequalities, among other economic, social, and environmental outcomes (Aschauer, 1989; Banerjee *et al.*, 2020; Baum-Snow *et al.*, 2017; Bird and Straub, 2020; Duranton *et al.,* 2014; Faber, 2014; Fedderke and Bogetic, 2009; Foster *et al.*, 2023a, 2023b; Michaels, 2008; Medeiros *et al.*, 2021a, 2022; Shi and Huang, 2014; Straub, 2011). Several related investigations have provided interpretable measures on the profitability of transportation investments by calculating economic return rates (RR), which have been used to guide cost-benefit analysis of infrastructure policies worldwide (Fernald, 1999; Li *et al.*, 2017; Medeiros *et al.,* 2021b; Wang *et al.*,2020).

Nonetheless, this strand of literature has focused on the efficiency goal of road policies. The efficiency aim is related to the fact that the road policy seeks to maximize economic returns by placing the infrastructure in regions with higher growth potential. This lucrativeness could be enhanced by allocating roads to localities that are more specialized in the transportation sector. The pioneering study by Fernald (1999) proposes a framework in which the national return rate to highway investments is calculated by multiplying a road-productivity elasticity parameter weighted by the national road reliance by an efficiency measure calculated by the Gross Domestic Product (GDP) divided by the national road stock. In a simple way, the greater the elasticity, the road specialization, and the efficiency, the greater the road investment return of a country, region, or locality.

While this empirical structure based on the efficiency-specialization aim has been followed by numerous researchers, a look at the redistribution and equity goals in the road investment return rates has been broadly overlooked. Redistribution is the policy purpose that uses road interventions to foster regionally balanced economic growth by targeting poorer localities. Similarly, equity means investing in places with low infrastructure endowments, equalizing the territory. In this context, some works tried to identify whether road policies are guided by efficiency or redistribution objectives and whether a trade-off exists between them (Cadot *et al.*, 2006; Fageda *et al.*, 2019; De la Fuente, 2004; Kemmerling and Stephan, 2002; Monastiriotis and Psycharis, 2014; Yamano and Ohkawara, 2000). Evidence is mixed, pointing out that road programs are guided by both efficiency and redistribution-equity sides and suggesting that there is some trade-off between the different policy goals. Although these studies have made significant contributions to the literature on infrastructure and development, identifying the road impact heterogeneities in terms of efficiency, road reliance, redistribution, and equity remains a *black box*. Our main contributions in this paper follow this path.

We investigate the highway investment impact heterogeneity on productivity in Brazil during the Growth Acceleration Program (PAC) period (2007-2018). There are three main reasons for making Brazil an interesting case study.

First, the PAC (divided into PAC 1 and 2) was the most important Brazilian infrastructure program in the last decades, doubling the level of investments in highway infrastructure compared to the previous ten years (Medeiros *et al.*, 2021b). At the same time, the program was criticized for its broad inefficiencies coming from harmful bureaucracy, unfinished buildings, and corruption, making the program's efficiency, efficacy, and effectiveness analysis critical and complex (Amann *et al.*, 2016; Burrier, 2019). Second, Brazil is characterized by deep regional inequalities in terms of both income and infrastructure endowment (Medeiros and Ribeiro, 2020). Third, the PAC period coincided (up to 2015) with relatively good economic performance in the Brazilian economy (Nassif *et al.*, 2020). Therefore, our case study is ideal for assessing the returns of infrastructure investments in a developing country scenario with marked regional disparities.

In addition, in August 2023, the Brazilian Federal Government launched the "New" PAC, predicting investments of around R$ 1.7 trillion in several infrastructure sectors. The main goals of the new (and third) program are similar to those of PACs 1 and 2, which are developing the precarious national infrastructure by augmenting public investments and attracting private resources to the sector. Then, an "Old" PAC evaluation is critical to identify bottlenecks and provide ways to upgrade the "New" PAC into a *win-win* policy, maximizing the economic returns of road investments considering redistribution and equity goals.

We estimate the impacts of national highway investments on local productivity growth at the municipality level using a long-difference econometric model (2007-2018). To overcome measurement error and non-random allocation biases in infrastructure studies, we combine Medeiros *et al.* (2024) third-step instrumental variable (IV) identification approach with the road reliance parameter by Fernald (1999). Next, we identify road impact heterogeneities regarding efficiency, road specialization, redistribution, and equity levels. It allows us to calculate an efficient-specialized and redistributive-equative return rate to highway investments in Brazil.

We find three main results. First, highway investment is highly productive in Brazil, and localities that are more dependent on roads benefit more from it. Second, the road impact on productivity is larger for less developed, poorly infrastructure-endowed, and more road-specialized municipalities. Third, road investment profitability appears to be deeply harmed when focusing on attending highly efficient places, likely related to huge infrastructure project costs and inefficiencies. In short, we provide evidence of win-win combinations for road policies in Brazil. In addition, targeting too efficient places and looking at their expected return might puzzle the inefficiencies related to them, overestimating the road investment impact on the economy. From those results, we calculate an average return rate to highway investments during the PAC of around 20%. This RR could have been larger by avoiding localities with very high efficiency levels and related inefficiencies and focusing on both efficiency-specialization and redistribution-equity goals, which does not seem to have occurred.

Our main contributions are fourfold. First, to the best of our knowledge, our study is the first one to include the redistribution-equity goals in the discussion on road investment profitability. We provide novel evidence by estimating road impact heterogeneities regarding both policy sides: efficiency-specialization and redistribution-equity. Second, our empirical approach allows us to identify "focal points" wherein road interventions might achieve *win-win* outcomes, i.e., profitability, redistribution, and equity. Third, we calculate and compare the PAC return rate with return rates considering different levels of efficiency, road reliance, redistribution, and equity, affording a unique ex-post evaluation of a national infrastructure program. Fourth, we contribute to the empirical literature on infrastructure by combining the third-step IV identification approach by Medeiros *et al.* (*2024*) with the infrastructure reliance framework.

This paper is structured as follows. Section 2 describes the related empirical literature. Section 3 presents the econometric approach. Section 4 details the data. Section 5 outlines the results. Section 6 evaluates the highway investment rentability in Brazil in light of the efficiency, road specialization, redistribution, and equity policy goals. Section 7 concludes.

3.2. Related literature

3.2.1. Road infrastructure and economic activity: empirical literature

A massive number of studies has investigated the relationship between infrastructure and economic activity (Baum-Snow *et al.*, 2020; Bird and Straub, 2020; Duranton *et al.*, 2014; Faber, 2014; Fedderke and Bogetic, 2009; Foster *et al.*, 2023a, 2023b; Jaworski and Kitchens, 2019; Straub, 2011). Since Aschauer (1989), several empirical investigations have provided evidence of the positive impact of transportation infrastructure investments on productivity and growth (Ghani *et al.*, 2014; Fahardi, 2015; Herzog, 2021; Holl, 2016; Li *et al.*, 2017; Zhang and Ji, 2019). In addition, road investments have been proven to alleviate regional and income inequalities (Medeiros and Ribeiro, 2020; Medeiros *et al.*, 2022) and poverty (Medeiros *et al.*, 2020; Parikh *et al.*, 2015), potentially constituting a *win-win* public policy.

Some papers have calculated the economic return rate to infrastructure investments to provide a clear and interpretable measure for policymakers. Fernald (1999) estimated a RR of 6% using United States data, Li *et al.* (2017) and Wang *et al.* (2020) found return rates for China of 11% and 23%, respectively, while Medeiros *et al.* (2021) and Medeiros *et al.* (*2024*) calculated return rates of 22.2% and 21.3%, respectively, using Brazilian data. In most cases, results point out that road investments are profitable, especially in the developing world scenario.

However, one of the main caveats of those papers concerns looking at only one side of public policy: efficiency. In other words, the calculated return rates are based on a constant road investment-productivity elasticity, and the national or regional road investment profitability will be higher the higher the ratio between the GDP and the road stock. Redistribution and equity public policy goals are set aside, and no heterogeneity in the road investment impact on productivity is investigated. For instance, we might expect more transforming effects of highway interventions in less developed and unconnected localities by expanding local markets and generating new activities and jobs (Jaworski and Kitchens, 2019; Storeygard, 2016). Disregarding these peculiarities might underestimate the return rates for underdeveloped localities and overestimate the profitability in developed economies. Next, we take into consideration other policy goals in road interventions.

3.2.2. Looking at efficiency-specialization, equity, and distribution in road investments

While efficiency has been widely studied in infrastructure studies, how road investments might heterogeneously impact local outcomes depending on redistribution and equity issues has remained overlooked. In this paper, we investigate the potential of transportation infrastructure policies to achieve *win-win* combinations characterized by regionally balanced economic growth.

We can identify four main features proposed by the related literature to identify the determinants of road allocation. Efficiency is understood as the profitability of road investments and is measured as the ratio between national, regional, or local GDP and road stock. Complementary, road specialization (or road infrastructure reliance) raises the impact of road investments on productivity by maximizing its return as the transport sector's role in the locality is more important. In other words, localities more specialized in road infrastructure benefit more from road investments. Those two first road features guide the principle of regional policy using the transport infrastructure to foster economic growth. On the other side, redistribution is mainly represented by GDP *per capita* or *per* worker, and the principle behind this variable is that the public policy is based on the use of transport infrastructure to promote the development of poorer regions. Equity is related to the idea of getting the territory right by means of road investments. This variable is measured as the total road stock over the geographical area. Then, road policies might have redistribution aims in terms of both income levels and infrastructure endowment.

A few studies have examined the existence of a trade-off between efficiency-specialization and redistribution-equity in road allocation (Albalate *et al.*, 2012; Cadot *et al.*, 2006; Fageda *et al.*, 2019; De la Fuente, 2004; Golden and Picci, 2008; Kemmerling and Stephan, 2002, 2008; Monastiriotis and Psycharis, 2014; Yamano and Ohkawara, 2000). In summary, this trade-off implies that, in most cases, investing in efficient-specialized, redistributive, and equative localities are conflicting aims. In other words, efficiency comes at the expense of redistribution and equity, and the opposite also occurs.

The findings are controversial. Some investigations found that governments intend to use road policies to promote economic growth, mainly guided by efficiency goals (Castells and Solé-Ollé, 2005; Kemmerling and Stephan, 2008), deepening regional inequalities (Cosci and Mira, 2018). On the other hand, other studies point out that transportation policies were guided by equative and redistributive aims (Fageda *et al.*, 2019; Fuente, 2004; Monastiriotis and Psycharis, 2014; Yamano and Ohkawara, 2000), and higher efficiency was expected to be reached if governments had chosen different targets.

We contribute to the infrastructure and regional development literature by evaluating the road investment impact heterogeneity in terms of the efficiency-specialization and redistribution-equity policy aims. Past studies examined how those road features have guided road allocation across countries and regions. However, our goal is to evaluate how these road characteristics affect the impact of road investment on productivity. From there, we provide novel evidence on the efficiency-redistribution trade-off and provide a practical way to identify places wherein *win-win* (efficient, road-specialized, redistributive, and equative) combinations could be achieved by investing in roads.

3.3. Econometric approach

3.3.1. Measuring causal impacts of highways investments on productivity: the Medeiros et al. (2024) three-step IV identification approach

To estimate the causal impacts of road investments on local economic development, we rely on the three-step IV identification approach proposed by Medeiros *et al.* (*2024*). The empirical approach deals with the endogeneity between highway investments and local outcomes in three sequential steps.

The first one concerns the study design. We evaluate the impact of a national road program (the PAC) implemented by the Brazilian Federal Government on outcomes at the municipal level. This study design is essential to alleviate endogeneity concerns between infrastructure investments and economic activity (Bird and Straub, 2020; Faber, 2014; Herzog, 2021), as we do not expect most municipalities to directly influence the Brazilian Federal Government decisions to place highways across the country.

The second step deals with measurement errors in the road measure. In the developing world scenario, infrastructure investments are likely impacted by inefficiencies such as corruption, harmful bureaucracy, and poor planning and execution. Whether it occurs, the relationship between road investments and economic development might be unclear, and biases are expected in conventional econometric estimates. Medeiros *et al.* (2024) tested several IVs related to the main geographical, environmental, and human physical infrastructure project costs to correct measurement errors in the road variable. They provided a range of suitable instruments, which we will replicate in this study. The rationality behind those IVs is that infrastructure costs might affect the outcome variable only through the highway variable (Holl, 2012; Lu *et al.*, 2022; Martín-Barroso, Nunez-Serrano, and Velazquez, 2015; Zhang, Hu, and Lin, 2020). For instance, we may expect local terrain ruggedness making a road project unfeasible or substantially more costly, which in turn is expected to affect a region or city economic development.

Nonetheless, it is unlikely this observable characteristic will directly impact GDP, population, or another economic outcome variable growth.

The third step corrects for the non-random placements of roads. For this, the authors draw several IVs based on hypothetical road networks trying to globally minimize road costs, political factors, and the propensity of municipalities to receive road interventions[15]. The hypothetical networks are constructed using the Least-Cost Path- Minimum Spanning Tree (LCP-MST) method (Dijkstra, 1959; Kruskal, 1956) following Faber (2014). The political instrument relies on Bird and Straub (2020), using the Brasília experiment as a source of quasi-random variation in road infrastructure construction. Finally, we use a third instrument based on traffic intensity data on federal roads, which Medeiros *et al.* (2024) named "potential road intervention areas." This IV deals with the fact that roads with heavy traffic are obvious candidates to receive interventions. In addition, we combine the IVs with the inconsequential unit approach by Chandra and Thompson (2000) by excluding central municipalities, which alleviates concerns about sample selection.

The econometric first (1) and second stage (2) equations are described as follows:

$$\begin{aligned} \Delta HighwayInvestments_{is} = {} & \beta_1 + \beta_2 CostRelatedIV_{is}^{suitable} + \\ & \beta_3 D_{is}^{New} \times NonRandomAllocationIV_{is}^{LCP-MST} + \\ & \beta_4 D_{is}^{New} \times NonRandomAllocationIV_{is}^{Political} + \\ & \beta_5 D_{is}^{Existing} \times NonRandomAllocationIV_{is}^{InterventionAreas} + \\ & \beta_6 X_{is} + \theta_s + \epsilon_{is} \end{aligned} \quad (1)$$

$$\Delta Y_{is} = \beta_7 + \alpha_1 \times \Delta HighwayInvestments_{is} + \beta_8 \times X'_{is} + \theta_s + \varepsilon_{is} \quad (2)$$

Where Δ represents the change between 2007 and 2018, Y_{is} is our dependent variable for municipality i in state s, $HighwayInvestments_{is}$ is the log form of the federal highway investments, α_1 is the elasticity of highway investments to our dependent variable, $CostRelatedIV_{is}^{suitable}$ is a vector containing suitable and tested cost-related IVs (Medeiros *et al.*, *2024*), $NonRandomAllocationIV_{is}$ is a vector of non-random allocation IVs, D_{is}^{New} is a dummy variable assuming value 1 whether a municipality were not connected by a federal road in 2006, and zero otherwise, $D_{is}^{Existing}$ is a dummy variable assuming value 1 whether a municipality were crossed by a federal road in 2006, and zero otherwise, X'_{is} is a vector of control variables, θ_s is a state fixed effect, β_1 to β_8 are parameters to be estimated, ε_{is} and ϵ_{is} are idiosyncratic error terms.

3.3.2. Including the local road infrastructure reliance (φ)

To investigate the road infrastructure reliance heterogeneity in Brazilian municipalities, we draw upon the studies by Fernald (1999), Li *et al.* (2017),

[15] For a more detailed description on the IVs construction, please see Medeiros *et al.* (*2024*).

Medeiros *et al.* (2021), Percoco (2015), and Wang *et al.* (2021) introducing a measure capturing the transportation infrastructure dependence of each municipality (φ). Then, our specification is slightly modified as follows:

$$\begin{aligned} \Delta HighwayInvestments_{is} \times \varphi_{is} = \beta_9 + \beta_{10} CostRelatedIV_{is}^{suitable} + \\ \beta_{11} D_{is}^{New} \times NonRandomAllocationIV_{is}^{LCP-MST} + \\ \beta_{12} D_{is}^{New} \times NonRandomAllocationIV_{is}^{Political} + \\ \beta_{13} D_{is}^{Existing} \times NonRandomAllocationIV_{is}^{InterventionAreas} + \\ \beta_{14} X_{is} + \theta_s + \epsilon_{is} \end{aligned} \quad (3)$$

$$\Delta Y_{is} = \beta_{15} + \alpha_2 \times \Delta HighwayInvestments_{is} \times \varphi_{is} + \beta_{16} \times X'_{is} + \theta_s + \varepsilon_{is} \quad (4)$$

Where β_9 to β_{16} are parameters to be estimated. By interacting the highway investment variable with the municipal road reliance, we control for local road specialization heterogeneity. When α_2 is positive, it implies that highway investments are productive, and municipalities depending more on road infrastructure benefit more from road investments. We contribute to the empirical literature on infrastructure by testing the three-step IV identification approach by Medeiros *et al.* (*2024*) combined with the infrastructure reliance framework (Fernald, 1999).

3.3.3. Calculating the return rate to highway investments: the efficiency-specialization, redistribution, and equity framework

Some studies have calculated the return rates to infrastructure investments around the world (Fernald, 1999, Li *et al.*, 2017, Medeiros *et al.,* 2021, Percoco, 2015; Wang *et al.*, 2021). Nonetheless, the literature has focused on the efficiency side of road investment following Fernald (1999). In this approach, the return rate to infrastructure investment is measured as follows:

$$RR = \alpha \times \varphi_{is} \times \frac{GDP_{is}}{Highway\ Stock_{is}} \quad (5)$$

Where RR is the return rate to highway investment, α is the road elasticity to productivity growth, and φ_{is} is the road reliance. The higher the elasticity and the road specialization, the higher the return rate. In addition, the return rate to highway investment will be larger the higher is the ratio between the GDP and the highway stock. Then, the efficiency (economic, measured as the GDP/Highway Stock ratio, and sectoral, measured as the infrastructure reliance parameter) guides the rationality behind the conventional return rates to road investments.

We contribute to the literature by recalculating the return rate to highway investments in the light of four interest variables: efficiency, redistribution, equity, and road specialization. To this end, we re-estimate Equations 3 and 4, slicing our sample into groups below and above the medians of those variables. We also try other specifications by excluding each variable's top 10% and 20% and the bottom 10% and 20% observations. Based on this approach, we provide novel results on the road investment impacts on productivity considering dramatic heterogeneities in the spatial distribution of efficiency, redistribution, equity, and road specialization. By doing this, we calculate several return rates considering the efficiency-equity variations in Brazil.

To better illustrate how our econometric exercises will be used to calculate heterogeneous RR to road investments based on the framework described in Section 2, we can rewrite Equation 5 as follows:

$$RR^{Sample_k} = \alpha^{Sample_k} * \varphi_{is} * \frac{GDP_{is}}{Highway\ Stock_{is}} \qquad (6)$$

Where *Sample* represents the different below and above the median of efficiency, redistribution, equity, and road specialization samples as described before, and k=1,...,53 is the number of different combinations between samples of our four interest variables. Then, we run one regression to each one of the proposed samples following Equations 1-4. For instance, we start by trying two samples (k=1,2) of municipalities below and above the median of efficiency. We repeat this procedure for the other three interest variables (k=3,...,8). Next, we combine samples of our four interest variables, generating a total of 53 different values for α[16]. By estimating different values for α, we can calculate different RR to highway investments weighting both goals of efficiency-specialization and redistribution-equity. If highway impacts on productivity are larger in less developed localities, our procedure allows us to correct the return rate by taking different values of α. We return to this application in more detail in the results and discussion section.

We re-run Equations 3 and 4 applying an additional econometric approach, including an interaction term between our highway investment variable and the efficiency, redistribution, and equity variables. We also interact the instruments with the same set of variables to allow the IV estimator to work.

3.4. Data

3.4.1. National highway investments

To measure the impact of national highway investments on local (municipal) economic activity, we use the unique dataset constructed by Medeiros *et al.* (2024)[17]. The authors georeferenced the PAC highway investments at the municipal level between 2007 and 2018 using data from the Ministry of Transportation and the National Highway System (SNV) georeferenced road data made available by the

[16] The possibilities are even more extensive. Nonetheless, we try to estimate combinations of samples considered “ideal samples”, which are represented by higher values of efficiency and road specialization and lower levels of redistribution and equity.

[17] An extensive description of the procedures needed to construct the dataset is provided by the authors.

National Highway Infrastructure Department (DNIT). Our main road variable is the sum of road investments (R$) in the whole period[18]. As a robustness check, we also try a dummy variable assuming value one if a municipality received a road intervention and zero otherwise.

3.4.2. Road efficiency, redistribution, equity, and specialization

We measure efficiency by the ratio between GDP and road stock by municipality. GDP can be extracted from the Brazilian Institute of Geography and Statistics (IBGE). To construct the road stock variable, we use Frischtak and Mourão (2017) sectoral estimates for the Brazilian road stock. The authors found a road stock of around R$ 594 billion in 2023 values. Next, we use georeferenced road data from the 2007 National Transport Logistics Plan (PNLT) to calculate the road length by municipality. We multiply single lanes by one and duplicated lanes by 2 to control for road quality and scale in our stock measure. Then, we divide the total road stock in monetary terms by our physical measure of road length to generate the monetary value by kilometer of road. Finally, we multiply this value by the road length of each municipality, which gives us our local road stock variable. Finally, efficiency is obtained by dividing the municipal GDP by its road stock.

Our redistribution variable is measured as GDP *per capita*. Both GDP and population are obtained from the IBGE. Equity is represented by the ratio between our calculated municipal road stock and the municipal geographic area (km^2).

Finally, our road specialization variable is calculated as the share of the municipal intermediate consumption related to the land transportation sector. This variable is constructed in three steps. First, we use wages and employment sector-level (four-digit CNAE classification) data from the Annual Social Information Report (RAIS) in 2007 to calculate the share of each sector in the total wages (employment). Second, we use the 2010 national Input-Output (I-0) table with 68 economic activities[19] to calculate our infrastructure reliance measure following Medeiros *et al.* (2021). Third, we aggregate the four-digit CNAE sectors into the sectoral classification of the I-O table. Then, we multiply the sector share in the total wages using the I-O sector-level classification by its respective infrastructure reliance value (φ_{sector}), obtaining the average infrastructure reliance by municipality, i.e., our road specialization measure (φ_{is}).

The spatial distribution of our four variables can be seen in Figure 3.1. Road investment efficiency is generally higher in the Southeast and South regions. A similar pattern can be observed for redistribution, indicating that the Southeast, South, and Mid-West regions have higher levels of GDP *per capita*. Equity is even more concentrated in the more developed regions of the country, also presenting

---

[18] Those investments include construction, paving, duplication, and enhancements. Maintenance expenses are not included due to data limitations.

[19] We exclude the land transportation sector to avoid obvious endogeneity issues, and the public administration sector as it has some practical issues as the concentration of public administration employment in central cities in the RAIS.

higher values in coastal areas. Regarding road specialization, we can observe a more decentralized distribution, showing that municipalities with different economic development levels and road endowments might be highly (low) specialized in road infrastructure.

Figure 3.2 reveals the allocation of road investments by deciles of efficiency, redistribution, equity, and road specialization. The Brazilian Federal Government focused on attending more efficient regions and more developed regions with a larger road endowment. For instance, an average municipality in the upper 10% of efficiency received approximately R$ 47.75 million, while an average municipality in the bottom 10% received around R$ 5.7 million. The same interpretation holds for equity and redistribution. Regarding road specialization, this characteristic does not appear to have driven the PAC road policy.

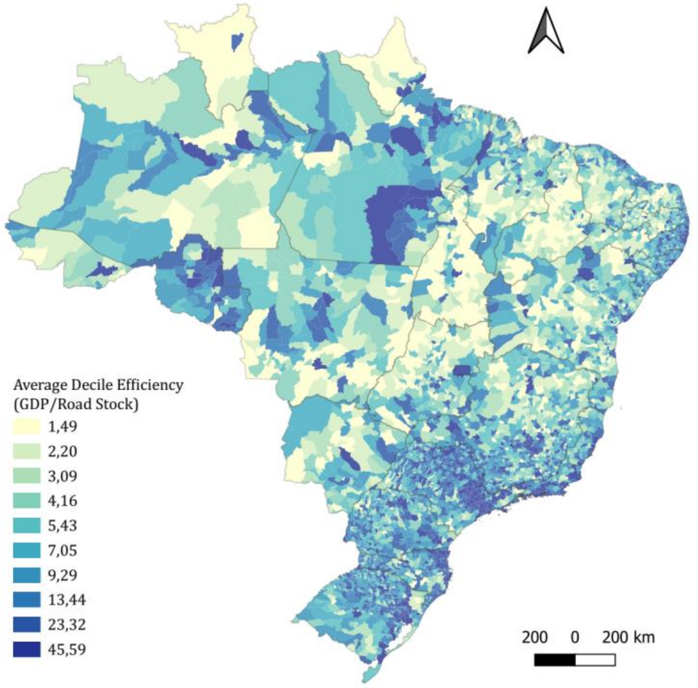
3.4.3. Instruments

We apply three main instruments in our econometric models using data from Medeiros *et al.* (*2024*). To correct measurement error in the road variable, we use two cost-related IVs (Cost Index 1 and Cost Index 2), which have been tested and considered suitable IVs. Both indexes were constructed using the Principal Component Analysis (PCA) by reducing the information of several variables representing the main infrastructure project costs. The first index represents environmental costs and is mainly composed of the share of legally protected areas and the number of environmental embargoes. The second index is more related to geographic costs, measured as the share of hilly areas, and human physical costs, measured as the share of urban infrastructure building. The higher the indexes, the higher the cost of constructing roads in the municipality.

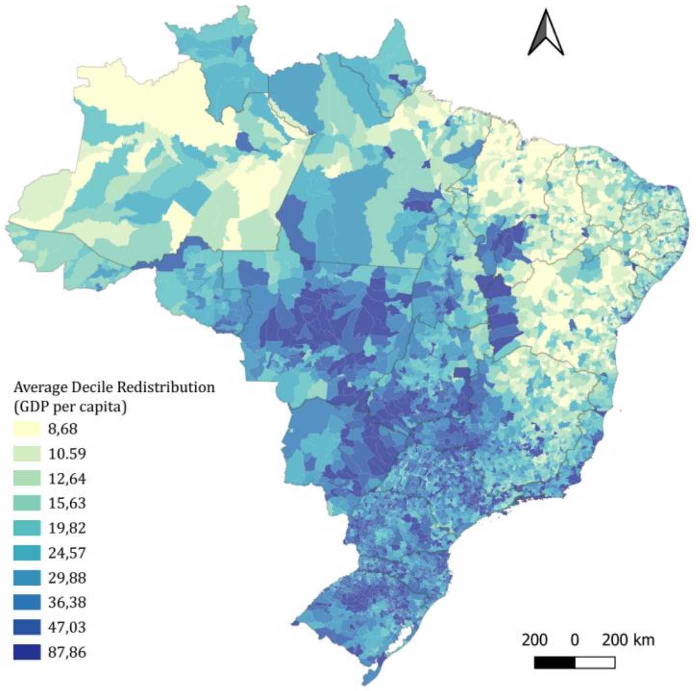
Our third instrument corrects for the non-random placement of roads. We use the Non-Random Allocation Index, which was created using the PCA technique. This IV represents the propensity of a municipality to receive a road intervention, as it is measured as the first principal component obtained from three original indicators. The first original indicator is the distance from a hypothetical network constructed using the LCP-MST method. In this case, a global minimization road network was generated, connecting the ending and starting points of those roads targeted by the PAC. The second original indicator is the distance from the hypothetical straight lines of the 1950s Brasília Plan following Bird and Straub (2020). The third original measure is the distance from a heavy traffic area. The higher the Non-Random Allocation Index, the smaller the propensity of a municipality to receive road investments. As robustness checks, we also try different combinations of instruments using the original indicators instead of the PCA indexes.

Figure 3.1. Road infrastructure in Brazil: efficiency (a), redistribution (b), equity (c) and road specialization (d)

(a)



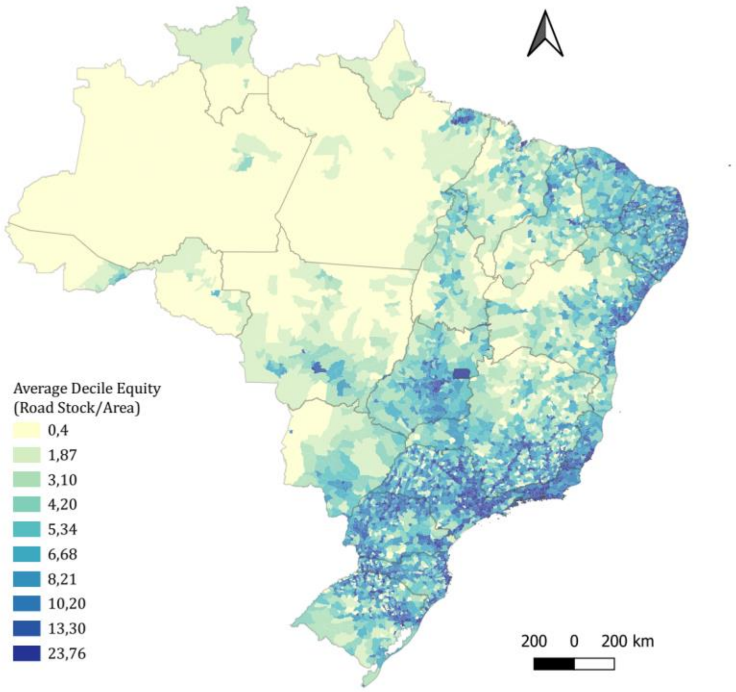
(b)



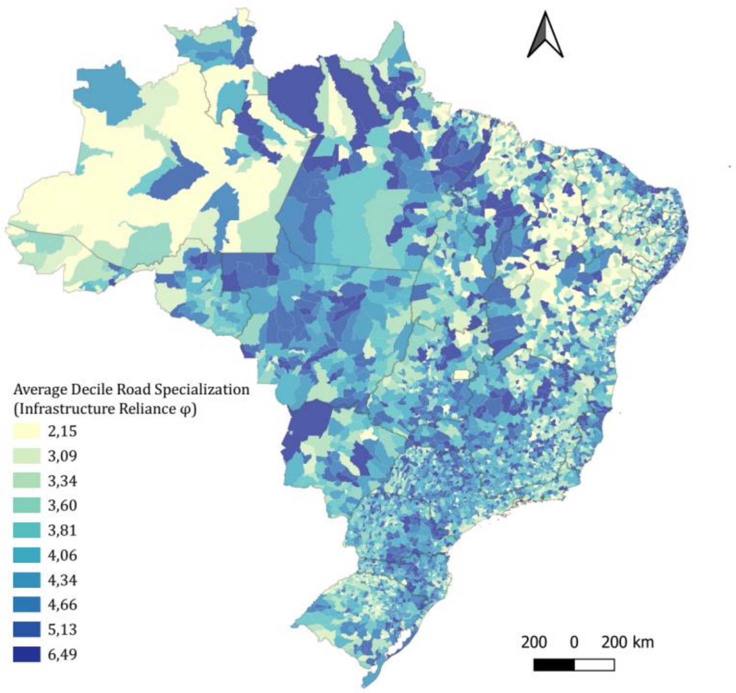
Source: authors' elaboration.

Figure 3.1. Road infrastructure in Brazil: efficiency (a), redistribution (b), equity (c) and road specialization (d)

(c)

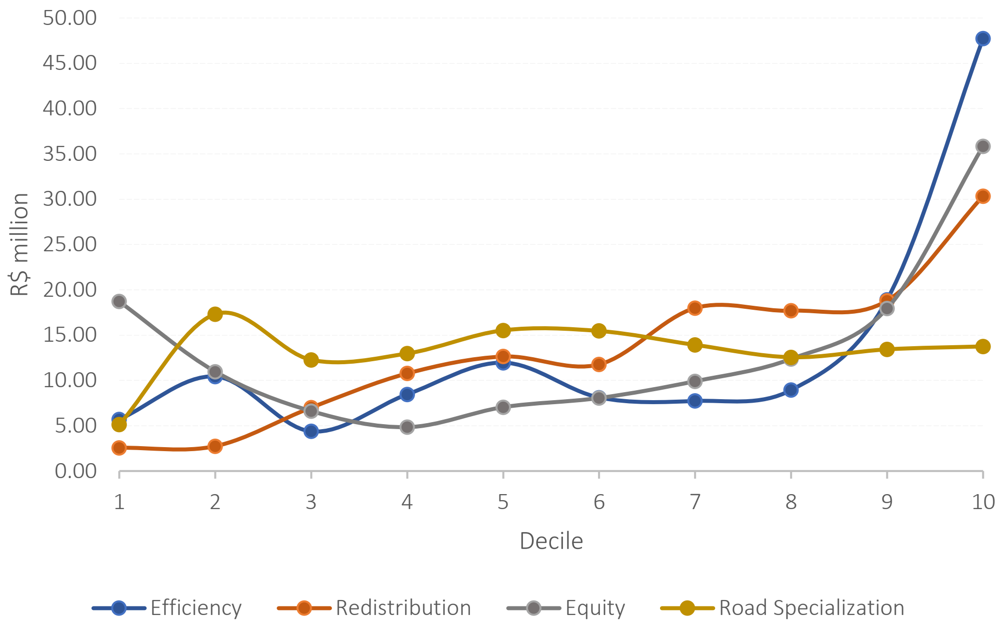


(d)



Source: authors' elaboration.

Figure 3.2. Average federal highway investments (R$ million) by deciles of efficiency, redistribution, equity, and road specialization



Source: authors' elaboration.

3.4.4. Dependent and control variables

Our primary dependent variable is GDP *per capita*. We use this variable as an extensively used *proxy* for municipal productivity, allowing us to calculate the return rates to highway investments. As robustness checks, we also try the number of workers (employment) and firms as dependent variables, data we get from RAIS.

To maintain consistency and comparability of results, we replicate the same set of control variables used by Medeiros *et al.* (*2024*). The vector of control variables is composed of state fixed effects dummies, the initial (2007) level of the dependent variable, the share of poor people, municipal geographic area, the number of workers, the agriculture share in the GDP, the share of exports in the national exports, the distance to the nearest state road, port and railroad, the number of railway stations in 1920, the distance to the capital Brasília, the Index of Municipal Institutional Quality (IQIM) and the population share with a master's or doctoral degree. By including this extensive set of variables, we control for several economic, social, institutional, geographical, and historical characteristics that could generate omitted variable bias in our regressions. A brief description of the variables used can be found in Table A1, and descriptive statistics can be seen in Table A2 in Appendix A.

3.5. Econometric results and discussion

3.5.1. Baseline estimates

Table 3.1 presents our baseline econometric results. In the first three columns, we replicate the same regressions by Medeiros *et al.* (2024)[20]. In columns 4-6, we include the infrastructure reliance parameter as specified in Equations 3 and 4. Cost-related and non-random allocation IVs are strong predictors of road investments and are quite suitable IVs, as demonstrated by the *Effective F* Statistics. Results are comparable to Columns 1-3 and remain unchanged for GDP per capita, employment, and the number of firms. Higher infrastructure project costs correlate with higher road investment levels, and the Non-Random Allocation Index negatively influences our highway variable, as expected.

Table 3.1. Federal Highway Investments and Local Outcomes Growth, 2007-2018: 2SLS IV Regressions

	1	2	3	4	5	6
	GDP *per capita*	Employment	Firms	GDP *per capita*	Employment	Firms
Second stage						
Log Highways Investments	0.0127***	0.0168***	0.0214***			
	(0.00)	(0.00)	(0.00)			
Log Highways Investments * φ				0.2928***	0.3733***	0.4781***
				(0.11)	(0.06)	(0.09)
First stage						
Cost Index 1	0.2190***	0.2207***	0.2376***	0.0084***	0.0084***	0.0086***
	(0.05)	(0.05)	(0.05)	(0.00)	(0.00)	(0.00)
Cost Index 2	0.3896***	0.3808***	0.4250***	0.0139***	0.0139***	0.0140***
	(0.07)	(0.07)	(0.07)	(0.00)	(0.00)	(0.00)
Non-Random Allocation Index	-0.5035***	-0.5069***	-0.5113***	-0.0216***	-0.0216***	-0.0219***
	(0.02)	(0.02)	(0.03)	(0.00)	(0.00)	(0.00)
Observations	5115	5141	5127	5115	5141	5127
Effective F Statistic	112.493	114.103	117.330	109.442	110.350	118.250
2SLS critical value for tau=5%	21.008	20.686	20.826	19.304	19.175	18.993
R^2	0.23	0.12	0.47	0.23	0.11	0.46

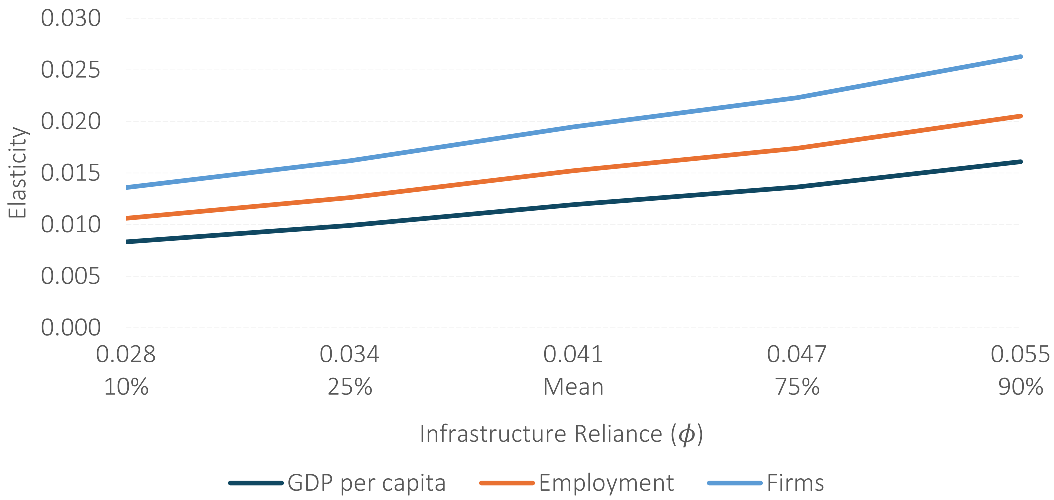
All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; workforce; agriculture share; exports share; distance to the nearest state road; distance to the nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors are reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Regarding the second stage regressions, we find positive parameters for the road investment variable. It implies that more road specialized municipalities benefited more from roads during 2007 and 2018. Those findings corroborate several other studies applying the infrastructure reliance approach (Fernald, 1999; Li *et al.*, 2017; Medeiros *et al.*, 2021; Percoco, 2015; Wang *et al.*, 2020).

[20] In the Medeiros *et al.*'s paper, the authors provide several robustness checks, including falsification tests for the IVs, different dependent variables, and alternative measures for the road investment. In unreported regressions, we run the same tests, and econometric results remain unchanged.

We can interpret our results using our road specialization parameter[21]. On average, we found a highway investment effect on GDP *per capita*, weighted by the municipal transportation infrastructure reliance, of around 0.29. To interpret this parameter in elasticity terms, we test several values of φ (Figure 3.3). Taking the average φ of 0.041, we get a road investment elasticity of around 0.012, slightly smaller than the average elasticity of 0.013 found by Medeiros *et al.* (2024), disregarding φ. Using the bottom 10% average φ, the elasticity is 0.008 while taking the upper 10% average φ, the elasticity is 0.016.

Figure 3.3. Highway investment elasticity by infrastructure reliance (φ) level



Source: authors' elaboration.

To better elucidate the importance of φ, we can reinterpret our results in terms of return rates. For this, we use Equation 5, supposing a constant GDP/Road Stock ratio of 16.7 (Frischtak and Mourão, 2017; Medeiros *et al.*, 2024). The average φ RR to highway investment is 20.04%, while the bottom 10% RR is 13.36%, and the upper 10% RR is 26.72%. In other words, the profitability of investing in the 10% more road-specialized localities is twice as much as the 10% less road-dependent. However, those findings do not explore the heterogeneities in the road effects on productivity depending on the local contexts of efficiency-specialization, redistribution, and equity. In the following regressions, we provide evidence on those issues.

3.5.2. Identifying heterogeneities in the return rate to highway investments

In this section, we investigate the role of efficiency, redistribution, equity, and road specialization on the allocation of roads and how those road features balance

[21] To test the significance of the local elasticities, we run nonlinear combinations of parameters tests (*nlcom* command in Stata) based on the delta method (Fieveson, 1999). All parameters shown in Figure 3 are significant at 1% level.

the highway investments impact on productivity. First, we test some correlations between our four interest variables and the highway investment measure to identify the Brazilian Federal Government's priorities in the placement of roads during the PAC. Next, we estimate several models by slicing our data below and above the median of our four interest variables, and we test interactions between our infrastructure characteristic variables and the highway investments. We also try some additional specifications as robustness checks.

3.5.2.1. Determinants of road investments

Table B1 in Appendix B shows the results of the determinants of road investment regressions. In Columns 1-5, we test our variables one by one and all together. As a robustness check, in Columns 6-10, we take the average values by decile of our four interest variables to avoid issues with outliers. Corroborating previous results showed in Figure 2, estimates suggest that efficiency was an essential factor in the placement of roads during the PAC. Similarly, the redistribution parameters are positive and significant in some regressions, indicating that municipalities with higher GDP *per capita* received more investments during the PAC. Regarding the road specialization variable, results suggest that more road-specialized localities received more highway investments. However, caution is needed as the infrastructure reliance parameter is significant at 5% or 10% significance levels in some specifications. There is weak and controversial evidence that equity was a relevant determinant of road allocation. Including equity without other characteristics, we find a negative correlation with road investments. Nonetheless, the signal becomes positive when we combine our road features, suggesting that localities with higher road endowments received more highway investments.

In short, the PAC appears to be mainly guided by efficiency and road specialization aims. In addition, the program invested more in wealthier localities, and there is weak evidence suggesting that the same pattern holds for localities with more extensive road endowment. These findings indicate a different priority in the road allocation in Brazil in comparison with other countries (Fageda *et al.*, 2019; Fuente, 2004; Monastiriotis and Psycharis, 2014; Yamano and Ohkawara, 2000), wherein a focus on equity and redistribution was found in contrast to efficiency. While those studies examined cases wherein there was strong road planning, as in Japan and European countries, our results put some caution on the PAC's role as the planner of road investments in Brazil in the recent past. By directing the road interventions to more prosperous and more profitable municipalities, the PAC seems to target places where a high economic return could be more straightforward achieved in the short run, neglecting the road investment potential to promote structural change and development in the less developed regions of the country (Medeiros *et al.*, 2020, 2022; Wang, 2022; Yang, 2018). In the next section, we explore whether investing in different municipalities with different levels of efficiency, road specialization, redistribution, and equity presents heterogeneous road impacts on productivity.

3.5.2.2. Road Investment Impact Heterogeneity

In this section, we provide novel results on the road impact heterogeneity in terms of efficiency, redistribution, equity, and road specialization. Past studies have examined these variables as determinants of road investment and allocation. In this paper, we are interested in identifying how those characteristics moderate the road impact on local productivity. For instance, we test whether localities below (above) the medians of our interest variables present different road-productivity elasticities. We expect road investments to be more profitable in more efficient and road-specialized localities following Equation 5, which was corroborated by our baseline econometric results. However, if the impact (α) of road investments is heterogeneous depending on the local levels of efficiency, road specialization, redistribution, and equity, governments might achieve win-win situations by designing efficient-specialized, redistributive, and equative road policies. Conversely, if there is a trade-off between efficiency-specialization and redistribution-equity, road policies will likely raise winners and losers across the country.

Table 3.2 presents the results of road investment heterogeneity. The first two lines exhibit the estimates by slicing our sample below and above the medians of our interest variables. Column 1 shows an unexpected result, demonstrating a higher road parameter for the bottom 50% of efficiency (0.71) than the upper 50% (0.28). Regarding redistribution, we found that roads significantly and positively impacted productivity only for the less developed municipalities (0.30). The road impact on GDP *per capita* growth also appears to hold for the more road-specialized localities (0.40). No heterogeneity effect was found concerning the equity goal.

Next, we try additional specifications by excluding the upper and bottom 10% and 20% of our four interest variables. These tests are important as we have a not negligible number of potential outliers. This is particularly relevant in our efficiency indicator. Brazil presents some municipalities with extremely high road efficiency values. The more extreme example is the city of São Paulo, presenting a GDP/Road Stock ratio of around 29,000 against a national ratio smaller than 17. This local ratio would imply a very high road profitability in the city. However, São Paulo presents huge infrastructure construction complexities related to human physical costs as expropriations and interferences because of its critical urban density and deep inefficiencies in road investments can be expected to occur there. The most critical example is the São Paulo *Rodoanel* (ring road), a sensitive infrastructure project started in 1998. Even today, the ring road is unfinished and presents a cost of R$ 34.4 billion, almost three times higher than its initially planned value. In addition, the city is the country's most populated and great economic center, which puts some caution on the role of road investments in those kinds of municipalities as expanding roads there might generate a constrained economic effect. Following the inconsequential unit approach, we dropped central cities in all our regressions, which alleviates this issue. However, outliers may have remained, so we proceed with the regressions in Lines 3 to 6.

Table 3.2. Federal Highway Investments controlling by infrastructure reliance (Log Highway Investments * φ) and GDP *per capita* Growth, 2007-2018: Heterogeneity Results, 2SLS IV Regressions

	1	2	3	4
Sample	Efficiency	Redistribution	Equity	Road Specialization
<=50%	0.7068***	0.3041***	0.2199	0.0356
	(0.15)	(0.11)	(0.17)	(0.23)
>50%	0.2814*	0.1016	0.2126	0.3944***
	(0.17)	(0.18)	(0.15)	(0.12)
>50% & <=90%	0.3328**	0.3261***	0.1019	0.4163***
	(0.17)	(0.11)	(0.20)	(0.14)
>50% & <=80%	0.6283***	0.1782*	-0.0130	0.4388***
	(0.18)	(0.09)	(0.24)	(0.16)
>10% & <=50%	0.6790***	0.2718***	0.5321**	0.2843
	(0.18)	(0.10)	(0.23)	(0.23)
>20% & <=50%	0.7214***	0.1946*	0.4764*	0.3321
	(0.19)	(0.10)	(0.29)	(0.25)

All regressions include the following set of control variables: GDP per capita in 2006; state fixed effects; municipality area; workforce; agriculture share; exports share; distance to the nearest state road; distance to the nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors are reported in parentheses. * 0.1 ** 0.05 *** 0.01.

By dropping the upper 10% and especially the upper 20% of efficiency, the road impact on productivity in the upper 50% converges to a similar level (0.63) compared to the bottom 50% (0.71). This result indicates a considerable inefficiency in allocating road investments to the most efficient municipalities, places that received the most PAC highway investments. The same pattern for redistribution suggests a similar road impact between the most developed (0.33) and the less developed (0.30) municipalities when dropping the upper 10% of redistribution. Regarding equity, excluding the bottom 10% to 20% of equity became the road impact on productivity positive and significant for the less road endowed localities. As we explore further, the environmental costs of investing in poorly road-endowed places might be too high to make road investment profitable in those regions. Finally, the road impact on GDP *per capita* only holds for highly road-specialized municipalities, remaining almost unchanged even when excluding the upper 10% and 20% of φ.

Those findings contribute to the specialized literature in several ways. First, our results suggest that there is not necessarily a trade-off between efficiency-specialization and redistribution-equity in Brazil. Employing Equation 5, the higher the efficiency, the elasticity (α), and the road reliance (φ), the higher the return rate to highway investment. Nonetheless, we find larger elasticities for those places characterized by lower levels of efficiency, economic development, and road endowment. In this sense, our results show a novel counterbalancing effect between investing in more efficient localities and taking a higher GDP/Road Stock ratio but a smaller α versus allocating roads to less efficient and developed municipalities and taking a smaller GDP/Road Stock ratio but a higher α.

A feature that seems to be an essential driver of road investment and its efficiency in our study, which has been overlooked by past studies, is the road specialization (φ) aim. Investing in more road-specialized places increases the return rate to highway investments utilizing the terms α (following Table 3.2 estimates) and φ. In addition, the road specialization indicator does not present a clear pattern of spatial concentration (Figure 3.1), being high or low in municipalities with high (low) efficiency, redistribution, and equity values. Then, φ increases the RR to road investments but does not necessarily generate a trade-off between efficiency and redistribution-equity. In the following econometric results, we try to better elucidate the role of road specialization in moderating the impact of roads on productivity.

In Table 3.3, we first slice our data below and above the median of road specialization. Then, we combine these samples with below and above the median samples of efficiency, redistribution, and equity. By doing this, we can identify whether the road impact on productivity persists for specific *win-win* samples, i.e., for samples of municipalities with high efficiency-road specialization, low GDP per capita, and wretched road endowment.

Column 1 presents the results for municipalities above the median of φ. In Line 1 of Column 1, we combine this above the median of φ sample with the above the median of efficiency municipalities. The same rationality holds for the subsequent lines, combining efficiency, redistribution, and equity samples. When combining the upper 50% of φ with the upper 50% of efficiency, the road impact on productivity is positive and significant. This impact is even higher when dropping the upper 10% and 20% of efficiency, corroborating previous estimates. The road impact is also favorable for those municipalities above the median of φ and below the median of redistribution. For equity, a positive and significant effect appears when dropping the bottom 10% of equity, reaffirming previous results.

In Columns 2 to 6, we combine samples above the median of φ with efficiency, redistribution, and equity samples. For instance, in Line 1 of Column 2, we test the road impact for municipalities above the median of φ, below the median of redistribution, and above the median of efficiency. In Column 2, results are not significant, suggesting that for those municipalities with higher road specialization, there is a trade-off between efficiency and redistribution goals.

In Columns 5 and 6, we re-estimate the specifications in Columns 3 and 4, respectively, but also select those municipalities below the median of redistribution. Results are weak, and caution is needed as our sample becomes relatively smaller, but the road impact on productivity remains positive and higher than the average parameter of 0.29 in Table 3.1.

Table 3.3. Federal Highway Investments, controlling by infrastructure reliance (Log Highway Investments * φ), and GDP *per capita* Growth by "Ideal Samples", 2007-2018: 2SLS IV Regressions

	1	2	3	4	5	6	7
Sample	Road Specialization > 50%	Road Specialization>50% & Redistribution <= 50%	Road Specialization > 50% & Equity <= 50%	Road Specialization > 50% & 10% <Equity <= 50%	Road Specialization > 50% & Redistribution <= 50% & Equity <= 50%	Road Specialization > 50% & Redistribution <= 50% & 10% < Equity <= 50%	Road Specialization <= 50%
Efficiency > 50%	0.5023***	0.2356	0.4172*	0.4823	0.3495	0.5071	-0.1897
	(0.16)	(0.16)	(0.23)	(0.34)	(0.27)	(0.35)	(0.41)
50%< Efficiency < 90%	0.5485***	0.2208	0.5444**	0.8612**	0.4139*	0.5188	-0.1756
	(0.17)	(0.17)	(0.24)	(0.39)	(0.25)	(0.38)	(0.39)
50%< Efficiency < 80%	0.6897***	0.2028	0.6259**	0.8161**	0.4452*	0.2728	0.1989
	(0.19)	(0.16)	(0.27)	(0.39)	(0.25)	(0.39)	(0.42)
Redistribution <= 50%	0.4169***	-	0.3218*	0.6186**	-	-	0.1492
	(0.12)		(0.18)	(0.26)			(0.18)
Equity <= 50%	0.2685	-	-	-	-	-	0.2146
	(0.18)						(0.34)
10% < Equity <= 50%	0.4114*	-	-	-	-	-	0.5596
	(0.22)						(0.48)

All regressions include the following set of control variables: GDP per capita in 2006; state fixed effects; municipality area; workforce; agriculture share; exports share; distance to the nearest state road; distance to the nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors are reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Finally, we run the same specifications of Column 1 but considering municipalities below the median of φ. In this case, we cannot observe significant road impacts on productivity, suggesting that higher values of road specialization are a critical condition to achieve road investment profitability in Brazil. As robustness checks, we run the same specifications in Tables 3 and 4, considering the road investment variable without the interaction term (φ). We present the estimates in Tables B2 and B3 in Appendix B. Results remain equal.

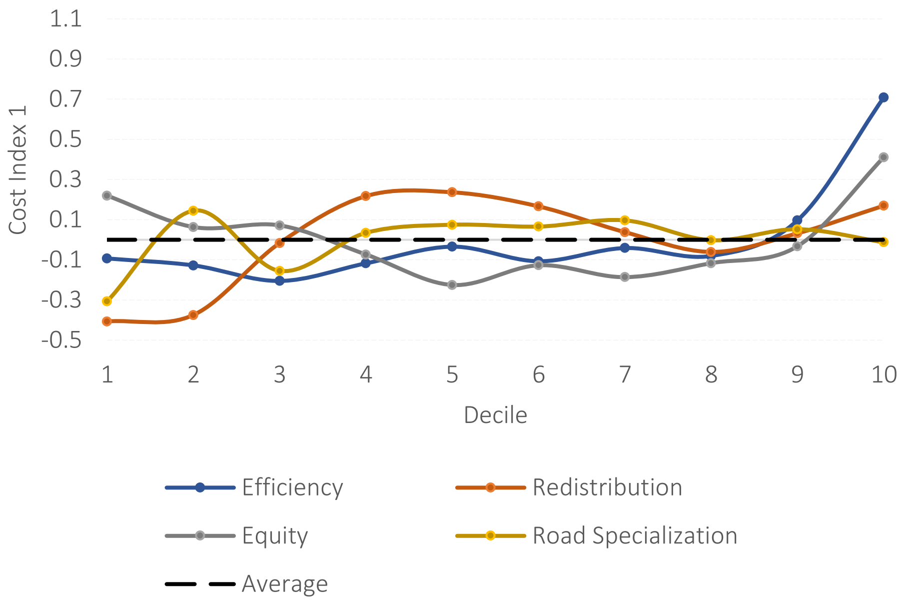
Next, we estimate the interaction models by multiplying the highway investment by our policy goal variables. Results are described in Table B4 in Appendix B. Results indicate that efficiency and redistribution present a significant moderating effect on the nexus between road investments and productivity. The road impact on GDP *per capita* falls as the efficiency and redistribution levels become larger. In other words, the effect of roads on the local economy is smaller for more developed and road-efficient places, corroborating previous estimates. Taking the 10% bottom values of efficiency and redistribution, a 1% increase in highway investments raises municipal productivity by 0.035% and 0.021%, respectively. Nonetheless, the elasticities are not significant when we take the 10% top values of efficiency and redistribution. In other words, the positive road impacts on productivity do not hold for very efficient and economically developed municipalities.

Now, we provide some explanations for our findings. Our results suggest that some road investment combinations could constitute *win-win* road policies in Brazil. In other words, we found positive, significant, and above-the-average road impacts on productivity, even considering both highly efficient-specialized and low redistribution-equity municipalities. However, we found a smaller road impact on productivity in more efficient localities compared to the less efficient ones. Interestingly, this impact is almost equalized when dropping municipalities in the upper 20% of efficiency. Those issues demand additional discussion.

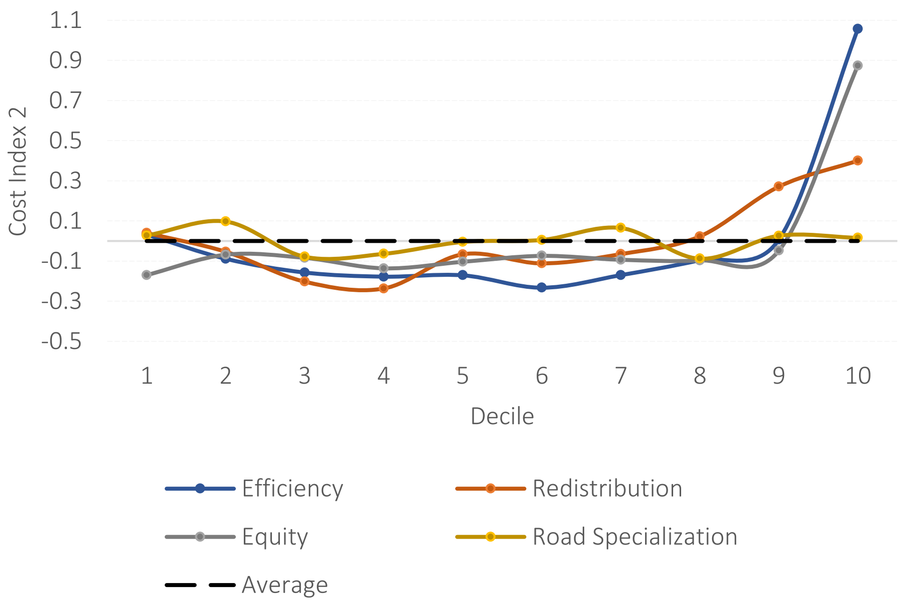
First, we might expect the road investment effects on GDP *per capita* to be larger in less developed regions in comparison with wealthier regions, as constructing new roads in isolated places may promote a deep process of structural change and economic development by expanding market integration and creating new activities and jobs (Jaworski and Kitchens, 2019; Storeygard, 2016). If efficiency is positively correlated with the levels of economic development, as it seems to be in the Brazilian case (Figure 1), this result is not unexpected. The evidence found in Table 3.3 indicates a larger road impact for the bottom 50% of efficiency and redistribution in comparison with its upper counterparts, which put some light in the same direction.

Second, the results obtained by excluding the municipalities in the upper 10% and 20% of efficiency seem strongly related to the Brazilian inefficiencies in the infrastructure sector. To better clarify this relationship, we calculate the Cost Index 1 and 2 average for each decile of our four interest variables (Figure 3.4).

Figure 3.4. Infrastructure project costs by decile: Cost Index 1 (a) and Cost Index 2 (b)



(a)



(b)

Source: authors' elaboration.

We can see costs relatively higher than the average for the municipalities in the upper 10% or 20% of efficiency, redistribution, and equity. The difference from the mean is more pronounced in the Cost Index 2, which represents geographic, expropriation, and interference costs. Regarding the Cost Index 1, which mainly captures environmental costs, there is a peculiar above-average value for the bottom decile of equity, representing higher environmental costs for the poorest connected places and might explain why, by excluding those municipalities, we get a positive and significant road elasticity for the bottom 50% of equity. These results shed some light on the role of broad inefficiencies in road allocation and how institutional shortcomings in the infrastructure sector during the PAC may have harmed road profitability across the country. Table B5 in Appendix B presents regression results by slicing our sample below and above the median of our two Cost Indexes. Our findings suggest that the positive road impact on productivity remains only for municipalities with smaller infrastructure project costs, corroborating our arguments.

3.6. Evaluating the economic return of road investments in Brazil during the PAC period (2007-2018)

On the one hand, we found some *win-win* combinations providing higher road investment elasticities for highly efficient-specialized and low redistribution-equity municipalities. On the other hand, the PAC allocated around R$ 37.1 billion to municipalities in the upper 20% of efficiency, representing more than 50% of the total PAC highway resources, excluding maintenance expenses. This allocation raises questions about the program's efficacy, efficiency, and effectiveness, as it focused on attending highly efficient, richer, and more costly localities wherein we have not found (or found smaller impacts compared with other samples) significant road impacts. The PAC's decisions and priorities played a significant role in the planning of the Brazilian highway infrastructure, and their impact should not be underestimated. In addition, by targeting the efficiency goal, the Brazilian federal government appeared to have served more short-term purposes, such as emergency buildings and heavy traffic areas, lacking what was expected from the PAC regarding its transforming role in the planning of the Brazilian highway infrastructure.

In addition, some period specificities might explain some of our econometric findings. First, Brazil has experienced deep economic, political, and institutional crises since 2015 (Arestis *et al.*, 2021; Nassif *et al.*, 2020). Second, the infrastructure sector institutional environment was poor during the PAC period, which contributed to several inefficiencies represented by a low budget execution, delays in road buildings, overpriced inputs, and corruption (Amann *et al.*, 2016; Armijo and Rhodes, 2017; Burrier, 2019; Raiser *et al.*, 2017). On top of that, the sharp drop in the public investment budget would further reduce the PAC's role as the country's infrastructure planner after 2015, with the Brazilian Federal Government focusing even more on attending urgent and critical road buildings.

In this context and considering the rough task of separating the planning and design issues of the PAC from the multiple crises Brazil faced during the period, we try to evaluate whether the PAC priorities in the road sector provided the most efficient, redistributive, and equative choice for the country. In other words, our primal aim is to answer whether the PAC could have achieved more satisfactory results (or at least the same outputs) in terms of the road investment profitability if the Brazilian Federal Government had opted for a more efficient-specialized, redistributive, and spatially fair road policy. Second, we provide novel return rate maps considering both goals of efficiency-specialization and redistribution-equity, allowing policymakers to identify focal points wherein *win-win* outcomes were expected to be achieved by investing in roads. This analysis opens up the possibility of more efficient, redistributive, and spatially fair road policies, offering a hopeful vision for the future of road investments in Brazil.

To this end, we calculate a range of return rates to highway investments following Equation 6, considering the road elasticity heterogeneities in Tables 3.3 and 3.4 and compare them with the PAC return rate. Table C1 in Appendix C summarizes the values of $\alpha^{sample}{}_{k}$ for each group of municipalities of our calculated return rates. The values of φ and the GDP/Road Stock vary by municipality for all return rates. As argued before, we take the average of each of our interest variables by decile to avoid issues with outliers. Then, we calculate the weighted average RR using $\alpha^{sample}{}_{k}$, φ, and GDP/Road Stock for the municipalities treated by the PAC and several other "ideal samples" considering groups of above the median of efficiency (excluding the upper 20% to avoid high costly localities), and above the median of road specialization, and below the median of redistribution, and below the median of equity (excluding the bottom 10% to avoid high environmentally costly places). We name those groups as "ideal samples" not just because they group road efficient-specialized municipalities and can be seen as a priority in economic terms, but above all because they assemble poorer and less road-endowed municipalities, constituting *win-win* samples whereby redistribution and equity goals are expected to be reached in a scenario of high economic profitability.

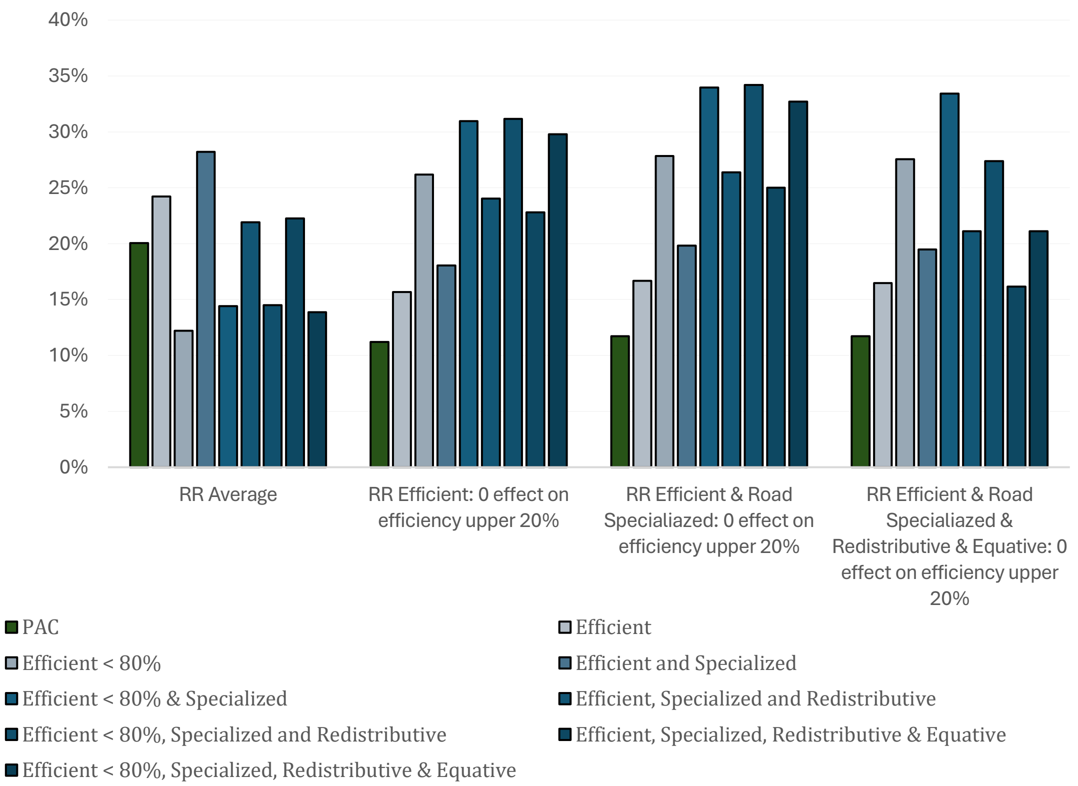
3.6.1. Was the "Old" PAC a win-win program?

Before starting our return rate analysis, we can compare the averages of our four interest variables in the PAC and our "ideal samples". Figure C1 in Appendix C shows the results. The PAC presented similar efficiency values compared to the sample of efficient (upper 50%) municipalities. The same pattern holds for redistribution. Thus, the PAC seems to have benefited more efficient and more affluent cities. Regarding equity, the municipalities impacted by the PAC showed a higher level of road endowment considering all the "ideal samples", suggesting that the PAC prioritized already road-connected regions even excluding maintenance expenses. By contrast, the municipalities receiving PAC investments presented one of the lowest levels of road specialization. These results suggest that the policy priority was mainly guided by efficiency by the side of GDP and not by road dependence, to some extent contraposing previous econometric findings.

Figure 3.5 summarizes the results in terms of return rates to highway investments. It is important to note that we are comparing the PAC with samples considered ideal because of both public policy aims of efficiency-specialization and redistribution-equity. In this sense, we do not compare the PAC with samples of less developed or road-endowed but not profitable municipalities, which would be unrealistic given the program goals.

We give examples to better elucidate how the return rates in Figure 3.5 are calculated following Equation 6. The "RR average" is our baseline return rate. It assumes α equal to 0.28 (Table 3.1) for all municipalities, allowing φ and GDP/Road Stock to vary by municipality. The PAC "RR average" (in green) is calculated by taking the weighted average of RR for the municipalities treated by the program, considering α equal to 0.28. The same rationality holds for the "ideal samples" (the additional eight columns in Figure 3.5). For instance, the "efficient" sample considers the weighted average of each proposed RR for those municipalities above the median of efficiency. Then, the "RR average" for the "efficient" group takes the average of RR for the municipalities above the median of efficiency, considering α constant (0.28) for all observations. Similar interpretations can be made for the other return rates, but in those cases, we also allow α to vary by groups of municipalities following our econometric estimates in Tables 3.2 and 3.3. For instance, in the "RR Efficient" return rate, we consider α equal to 0.71 for municipalities below the median of efficiency, 0.63 for those above the median but below the upper 20%, and zero for the remaining observations. The more restrictions we place on our interest variables, the smaller the number of municipalities treated as "ideal samples" and the smaller the external validity of our identification strategy.

Figure 3.5. Return rate to highway investments in Brazil: looking at heterogeneities in efficiency, redistribution, equity, and road specialization



Source: authors' elaboration.

The first three return rates ("RR Average", "RR Efficient" and "RR Efficient & Road Specialized") focus on the efficiency-specialization goal, as we do not impose any variation in α in terms of redistribution or equity. For those return rates, the "ideal samples" characterized by being efficient and road specialized naturally present the larger returns. The fourth rate ("RR Efficient & Road Specialized & Redistributive & Equative") includes redistribution and equity in the discussion.

As expected, the PAC presents a comparable "RR Average" value concerning the "ideal samples", reaffirming its efficiency goal. However, when we allow α to vary, other samples become more profitable than the PAC one, suggesting a lack of priority of the program in terms of road specialization, redistribution, and equity, and its likely excessive focus on more costly and developed localities. Interestingly, the return rates for our "ideal samples", even when considering redistribution and equity goals, are comparable to those related to the efficient and road-specialized municipalities and, in most cases, higher than the PAC return rate. This result arises from the higher α for less developed places that more than compensates for the dropping in the GDP/Road Stock ratio when evaluating our "ideal samples." For instance, the PAC "RR Efficient & Road Specialized & Redistributive & Equative" return rate, which allows α to vary by groups of all our four interest variables, is 11.7%. In contrast, the same return rate for the efficient-specialized and below the median of redistribution and equity municipalities is 16.1%, reaching 21.1% when we drop the 20% more efficient municipalities. A more profound redistributive and equative road policy between 2007 and 2018 could have achieved even higher economic returns.

It is worth noting that we are supposing a zero effect (α) for the municipalities in the upper 20% of efficiency. Based on regressions in Tables 3 and 4, investments in the upper 20% of efficiency harm the road profitability in the country likely because of the high infrastructure project costs and the related institutional and sectoral fragilities. In addition, we might expect a larger and more transforming road impact on productivity in less developed and road-endowed municipalities. The PAC return rate drops considerably from 20% to around 12% by supposing the zero effect. The same pattern holds for the fully efficient samples. Note that the PAC return rate under the zero effect becomes the smallest one, reinforcing that vast investments were destined for more prosperous places where higher economic returns were foreseen. On the other hand, the return rates of the samples excluding the upper 20% of efficiency are naturally less affected. In these cases, we observe return rates to highway investment from 21.1% to 34.2%.

Our findings can be interpreted in the light of two additional channels. First, the Brazilian infrastructure sector inefficiencies during the PAC seemed deeply harmful (Amann *et al.*, 2016; Armijo and Rhodes, 2017; Burrier, 2019). It is likely that hugely investing in highly costly localities collided with environmental licenses, expropriation and interference disputes, land conflicts, and other issues translated into schedule delays, unpredicted financial resources, and unfinished buildings. Complementary, the Brazilian institutional, political, economic, and social deterioration from 2015 imposes sharp drops in the PAC investments, which contributed to the stoppage of several interventions and their efficacy. Second, focusing on the efficiency goal does not seem to be the best option, even in terms of profitability. By not looking at the heterogeneities in the road impacts on productivity, in which we found significant effects for efficient and specialized regions but also for lower road-endowed and with smaller GDP *per capita* localities, the country invested in places wherein the return rate was lower than it could have been. In other words, the road investment policy in Brazil during the 2007-2018 period could have been two to three times higher by targeting efficiency-specialized and redistribution-equity places and overcoming infrastructure sector inefficiencies.

3.6.2. An efficient-specialized and redistributive-equative return rate map

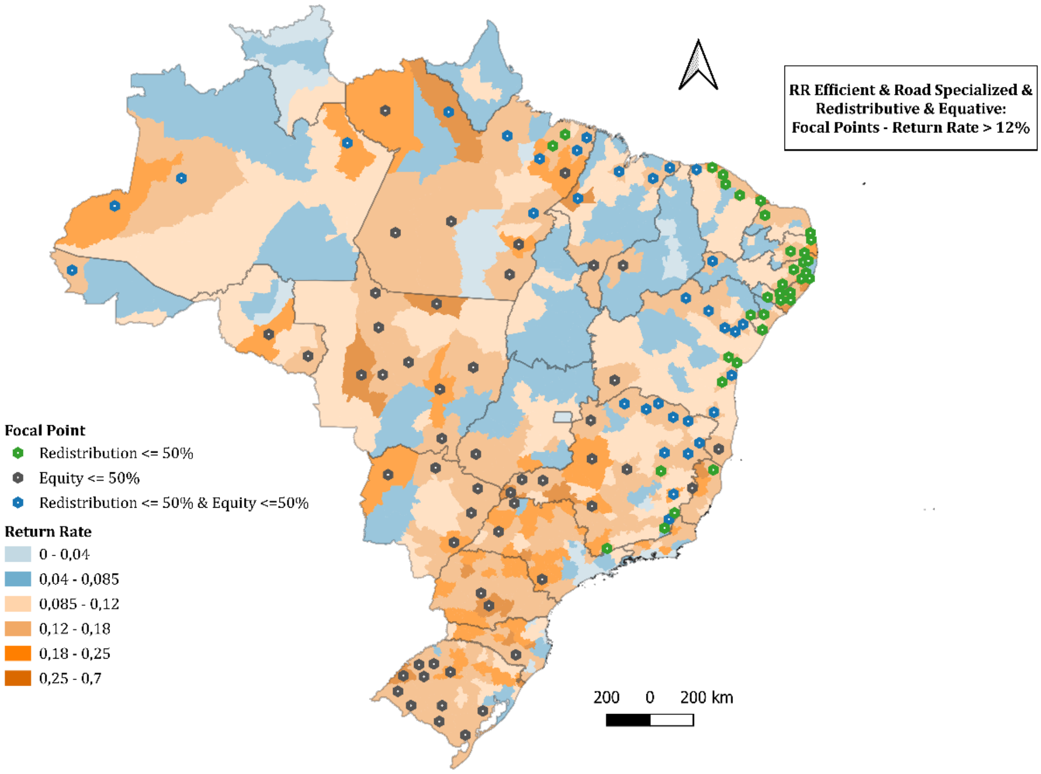
This section provides a return rate map based on our estimates considering the road impact heterogeneities. We do not expect the Brazilian Federal Government to target specific municipalities in allocating roads (Medeiros *et al.*, 2024), so we construct the maps at the Immediate Geographical Region (RGI) level. The 510 RGIs are groups of municipalities in the urban network sharing a common local urban center as their basis, being constructed by the IBGE. Its design considers the connection of nearby cities through dependency relationships and the population's movement in search of goods, services, and employment opportunities. Then, the RGIs are closely related to transportation goals and can be seen as a reasonable spatial scale in terms of highway public policies. The RGI level return rates are calculated using the municipal averages of the GDP/Road Stock ratio and φ. Then, the RGI measure is an input for evaluating the efficiency goals on both sides of GDP and road specialization.

Next, we create three groups of focal points, i.e., regions representing *win-win* policy intervention combinations, as follows: i) municipal redistribution average below the median; ii) municipal equity average below the median; and iii) municipal redistribution and equity average below the median. We consider a region minimally profitable if its RR is higher than the Social Discount Rate (TSD) of 8.5% (Ministry of Economy, 2021). However, the Brazilian basic interest rate is above 12% nowadays, which puts some concern on the profitability levels between 8.5% and 12%. Then, we filter those focal points with RR above 12% to guarantee a reasonable cut-off.

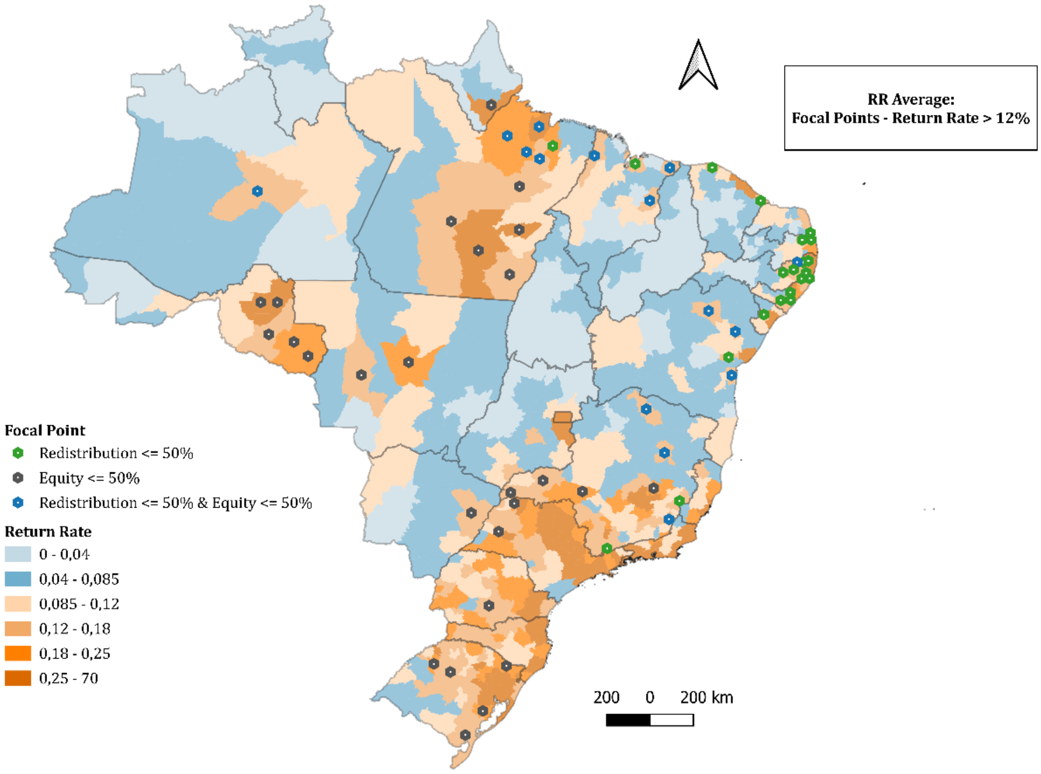
Figure 3.6 exhibits the return rate maps. Figure 3.6 (a) is our preferred return rate, which allows α to vary by municipality groups of efficiency, road specialization, redistribution, and equity ("RR Efficient, Road Specialized, Redistributive, and Equative" in Figure 5). Figure 3.6 (b) shows the "RR Average" measure, which we use as the RR of comparison as heterogeneities in α are not considered.

Figure 3.6. RR Efficient & Road Specialized & Redistributive & Equative (a) and RR Average (b): Pointing *Win-Win* Localities for Road Investment

(a)



(b)



Source: authors' elaboration.

As we can see in Figure 3.6, in the "RR Average" measure, the profitable localities are concentrated in the Southeast and South regions, coinciding with the most developed part of the country. On the other hand, when we allow α to vary by levels of redistribution and equity, a new range of profitable areas is open, especially in the inland and poorer regions. Significantly, the number of focal points increases, suggesting that a short-sighted look into the efficiency-specialization goals through the conventional RR measure might puzzle the Brazilian road investment profitability heterogeneity.

We can decompose our return rate into economic and equality components. To obtain the equality component, we subtract the RR Efficient & Road Specialized & Redistributive & Equative from the RR Average. This difference shows us the increase (decrease) in the return rate to highway investments due to redistributive and equative issues. Therefore, the economic component can be taken by the difference between the RR Average and the equality component. Figure C2 in Appendix C exhibits the economic and equality components. As argued before, higher economic returns are concentrated in more developed localities in the South and Southeast and in some areas where the agriculture sector has boosted the Brazilian economy. On the other hand, the equality component is higher in poorer and more isolated regions. It is interesting to note that the equality component is larger than the TSD for some RGIs, suggesting that the dropping in road-related inequalities more than compensates for investing in roads there, even ignoring the economic component.

Regarding our focal points, we can identify some patterns. First, the equity focal points (in gray) are spatially concentrated in the Mid-West and South regions and the south of Pará. Those regions are characterized by a suitable level of GDP *per capita* but a low level of road endowment, suggesting that they are places with substantial growth potential and high demand for highway investments. Second, the redistribution focal points (in green) are localized in the Northeast region, places marked by low productivity levels but a proper road provision. Finally, redistribution and equity focal points (in blue) are mainly present in the North and Northeast regions and in the north of the State of Minas Gerais, places distinguished by being economically and road-endowed poor. It is worth noting that most of the RR map is orange (above 8.5%), implying that Brazil needs considerable road investments, as we expected, given the lame national and regional infrastructure sector scenario.

3.7. Concluding remarks

Using a novel empirical approach, we estimated the impact of road investments on productivity in Brazilian municipalities between 2007 and 2018. More specifically, we evaluated how heterogeneities in road efficiency, specialization, redistribution, and equity affect the return rate to highway investments across the country.

Our main results pointed out a positive impact of road investments on productivity, corroborating several past studies. In addition, we found higher road impacts on GDP *per capita* for less developed and road-endowed places and lower impacts for highly efficient municipalities. Those findings suggest the existence of *win-win* combinations for road policies in Brazil, in which policymakers might maximize the road economic returns by investing in both efficient-specialized and less developed localities. In addition, huge infrastructure project costs and institutional and sectoral inefficiencies likely harm the predicted high profitability in extremely efficient places.

While we contribute to the empirical literature on infrastructure and development, some open points remain. First, we cannot affirm whether our results hold for other infrastructure sectors even during the same period. Second, we compared the PAC with samples considered ideals in the sense that they represented efficient-specialized and redistributive-equative groups. However, infrastructure might promote a deep process of structural transformation and development in less developed and unconnected regions, and a more careful look into the redistributive-equative municipalities disregarding efficiency goals is needed. Third, road investments are expected to foster economic and social prosperity, but these gains may be accompanied by environmental damaging as increasing deforestation and greenhouse gas emissions. We hope to provide some new evidence on those issues in future research.

4. HIGHWAY INFRASTRUCTURE AND GREENHOUSE GAS EMISSIONS:
evaluating the environmental costs of road investments in Brazil

Abstract

This study evaluates the impacts of highway infrastructure development on greenhouse gas (GHG) emissions. To this end, we use detailed local-level data from Brazilian municipalities during the Growth Acceleration Program (PAC) period (2007-2018) and apply an instrumental variable identification approach to circumvent endogeneity concerns related to the non-random placement of roads. We find that a 1% increase in road investments raises CO2 emissions by 0.025%. Those damaging highway effects are sustained for the road, energy, and land use change sectors. In addition, findings point out heterogeneous road impacts on CO2 emissions depending on agglomeration, population scale, deforestation, and technology. From the econometric estimates, we calculate an average CO2 emissions return rate to highway investments (ERR) of 3.0%, implying a discount on the economic benefits of road investments proved in past studies. Finally, we measure a sustainable return rate to highway investments (SRR) of around 17%, indicating a widespread need to develop the Brazilian transportation sector. It is important to note the deep regional heterogeneities in Brazil, wherein we can observe negative SRRs for some regions. This research offers valuable insights for policymakers, technicians, financial institutions, and civil society in shaping effective and environmentally conscious road policies.

Keywords: transportation infrastructure; regional development; sustainability; greenhouse gas emissions.

4.1. Introduction

A broad strand of literature has proven the positive role of transportation infrastructure on economic growth and productivity (Aschauer, 1989; Baum-Snow *et al.*, 2020; Bird and Straub, 2020; Faber, 2014; Foster *et al.*, 2023a, 2023b; Ghani *et al.,* 2014; Herzog, 2021; Jaworskiy and Kitchensz, 2019; Straub, 2011; Zhang e Ji, 2019). Some of those investigations have calculated economic return rates to highway investments as a measure of its profitability, which are used to guide cost-benefit analysis and transportation policies around the world (Fernald, 1999; Li *et al.*, 2017; Medeiros *et al.,* 2021b; Medeiros *et al.*, 2024; Wang *et al.*,2020). While those studies have provided important results, the environmental costs (or benefits) of highway investments are put aside (Alam *et al.*, 2022; Quadros and Nassi, 2015; Laird and Venables, 2017; Welde and Tveter, 2022).

In this paper, we evaluate the unclear relationship between highway infrastructure development and greenhouse gas (GHG) emissions. On the one hand, road construction and enhancement tend to increase GHG emissions in the construction and maintenance phases by the direct use of materials and equipment. Once the highway is built, the growth in the road network increases regional accessibility, population mobility, and interregional traffic flows, boosting transportation demand and affecting GHG emissions. On the other hand, road development might decrease GHG emissions by reducing travel time and distance, which lowers GHG emissions during transportation, as well as by stimulating agglomeration economies, reducing energy consumption, and boosting energy efficiency. Empirical findings are mixed, pointing out increasing effects (Churchill *et al.*, 2021; Emodi *et al.*, 2022; Ghannouchi *et al.*, 2023; Lin *et al.*, 2017; Luo *et al.*, 2018; Xiao *et al.*, 2023; Xie *et al.*, 2017; Yao *et al.*, 2023) as well as null or reducing impacts (Georgatzi *et al.*, 2020; Ghannouchi *et al.*, 2023; Han *et al.*, 2017; Li and Lu, 2022) of highway infrastructure on GHG emissions. In addition, there are heterogenous road impacts on GHG emissions depending on agglomeration, development level, economic growth, population scale, and technology, among other moderating variables.

While the literature on road infrastructure and GHG emissions has provided relevant findings and discussed critical transmission channels, some gaps remain. First, to the best of our knowledge, there are no studies that calculate a sustainable return rate to highway investments, i.e., adding (discounting) the environmental benefits (damages) from the broad evaluated economic returns of road investments, which would be relevant to infrastructure policy planning, design, financing, and evaluation. Second, investigations using detailed local-level data are scarce, and the existing literature relates to China (Han *et al.*, 2017; Li and Lu, 2022; Luo *et al.*, 2018; Xiao *et al.*, 2023; Xie *et al.*, 2017; Yao *et al.*, 2023). Using local municipal (city) data might capture important heterogeneities across the space, providing new evidence to the specialized literature. Third, most papers have studied European countries or China, wherein the energy and industry sectors are the most important sources of

GHG emissions. Then, analyzing cases in which other sectors, such as land use change and agriculture, are more relevant to GHG emissions might shed some light on new transmission channels and heterogeneous impacts of road development on the environment. We seek to contribute to the literature in those directions.

We evaluate the impact of highway investment on GHG emissions growth in Brazilian municipalities during the Growth Acceleration Program (PAC) period (2007-2018). To this end, we use detailed local-level data on national road investments and GHG emissions and apply an econometric approach dealing with the endogeneity coming from the non-random placement of roads, allowing us to identify causal road impacts on greenhouse gas emissions. From these estimates, we calculate carbon dioxide equivalent emissions return rates (ERR) and sustainable return rates (SRR) to highway investments in several Brazilian localities.

The Brazilian case study is interesting for several reasons. First, the PAC (divided into PAC 1 and 2) was the most important Brazilian infrastructure program in the last decades, doubling the level of investments in highway infrastructure compared to the previous ten years (Medeiros *et al.*, 2021b). Second, Brazil presents deep regional heterogeneities in terms of infrastructure endowment, income (Medeiros *et al.*, 2021a, 2022; Medeiros and Ribeiro, 2020), and GHG emissions. Third, unlike most studies evaluating the Chinese and European cases wherein energy and industry sectors are the most important sources of GHG emissions, the Brazilian economy presents the land use change and agriculture sectors as the main contributors to GHG emissions. Our case study is ideal for evaluating road investments' economic and environmental profitability in a developing country context with huge regional disparities and for providing novel transmission channels from roads to the environment in a unique environmental scenario.

Furthermore, the Brazilian Federal Government launched the third PAC in August 2023. To the best of our knowledge, this is the first time in Brazilian history that an extensive national infrastructure program has included explicit environmental proposals. As one of the main mechanisms to foster environmental practices in the infrastructure sector, the Brazilian Government prioritizes and facilitates the availability of funds to projects with environmental devices promoting and accelerating the ecological transition. In the transportation sector, the “new” PAC presents the "Efficient and Sustainable Transport" pillar, which deliberates investments of around R$ 349.1 billion in several transportation buildings, including the road sector. Additionally, the program provides several institutional initiatives related to environmentally suitable road infrastructure. For instance, the program incentivizes the ecological transition by issuing sustainable sovereign bonds, expanding the Climate Fund (*Fundo Clima*) resources, promoting low-carbon transportation such as hybrid and electric vehicles, and encouraging decarbonization and using sustainable materials in the construction sector.

While those policy tools are critical to Brazilian sustainable development, a precise regionalized measure of highway investments' environmental costs (or benefits) is lacking. In this context, evaluating the “old” PAC – in which emphatic environmental initiatives related to the road sector were mostly absent– is critical to provide evidence on the environmental costs of road investments, maximizing its economic returns while respecting environmental preservation and recovery. Then, a novel measure of sustainable return rate to highway investments might represent a key input to policymakers, technicians, financial institutions, and civil society in planning, designing, financing, and evaluating current and future road policies.

In this context, we find three main results. First, we find that a 1% increase in road investments raises GHG emissions by 0.025%. This result is maintained under various specifications capturing heterogeneous road impacts and several robustness checks. Second, we calculate an average GHG emissions return rate to highway investment (ERR) of 3.0%, demonstrating a harmful environmental impact of roads. By subtracting our ERR from the economic return rate to highway investments (RR) from Medeiros *et al.* (forthcoming), we find an average Sustainable Return Rate to Highway Investments (SRR) of around 17%, indicating a widespread need to develop the Brazilian transportation sector even considering its environmental costs. To reduce our average SRR of 17% to the threshold of 8.5%, Brazil would need two times more highways, which implies a road stock of 14% of national GDP, in line with Frischtak and Mourão (2017) and Medeiros *et al.* (2021b). Third, we find critical regional heterogeneities in our ERR and SRR. In general, the environmental damage from roads is more pronounced in less populated and poorer localities, which coincides with some critical areas in the Brazilian Amazon.

Our main contributions to the specialized infrastructure and regional development literature are fourfold. First, we propose two novel regional measures related to the environmental costs of highway investments: i) the CO2 Emissions Return (Discount) Rate to Highway Investments (ERR) and ii) the Sustainable Return Rate to Highway Investments (SRR). In doing so, we provide novel, easy-to-interpret measures in the context of political decision-making. Second, we provide original evidence on the relationship between highway infrastructure development and GHG emissions in a context wherein land use change and agriculture sectors are the most critical contributors to GHG emissions. Third, we advance in relation to past studies by evaluating new heterogeneous road impacts on GHG emissions, mainly related to the environmental and institutional weaknesses from deforestation and illegal land use. Fourth, we circumvent endogeneity issues from the non-random placement of roads by adapting an instrumental variable identification approach to the GHG emissions context.

This paper is structured as follows. Section 2 describes the related empirical literature. Section 3 presents the methods and data. Section 4 outlines the econometric results. Section 5 provides the results regarding the sustainable return rate to highway investments. Section 6 concludes.

4.2. Related Literature

4.2.1. Transportation infrastructure and economic development

A massive strand of literature has investigated the relationship between transportation development and economic activity (Baum-Snow *et al.*, 2020; Bird and Straub, 2020; Duranton *et al.*, 2014; Faber, 2014; Fedderke and Bogetic, 2009; Foster *et al.*, 2023a, 2023b; Jaworski and Kitchens, 2019; Straub, 2011). Since the pioneering study by Aschauer (1989), several empirical studies have proven a positive role of highway investments on productivity and growth (Ghani *et al.*, 2014; Fahardi, 2015; Herzog, 2021; Holl, 2016; Li *et al.*, 2017; Zhang and Ji, 2019).

Related papers calculated economic return rates (RR) to infrastructure investments to provide an easy-to-interpret measure for policymakers and the society. Fernald (1999) measured a RR of 6.0% using United States data, Li *et al.* (2017) and Wang *et al.* (2020) found return rates for China of around 11% and 23%, respectively, while Medeiros *et al.* (2021, *forthcoming*) and Medeiros *et al.* (2024) measured RRs around between 20% and 22.2% using Brazilian data. In general, findings confirm that road investments are profitable, especially in the context of the developing world.

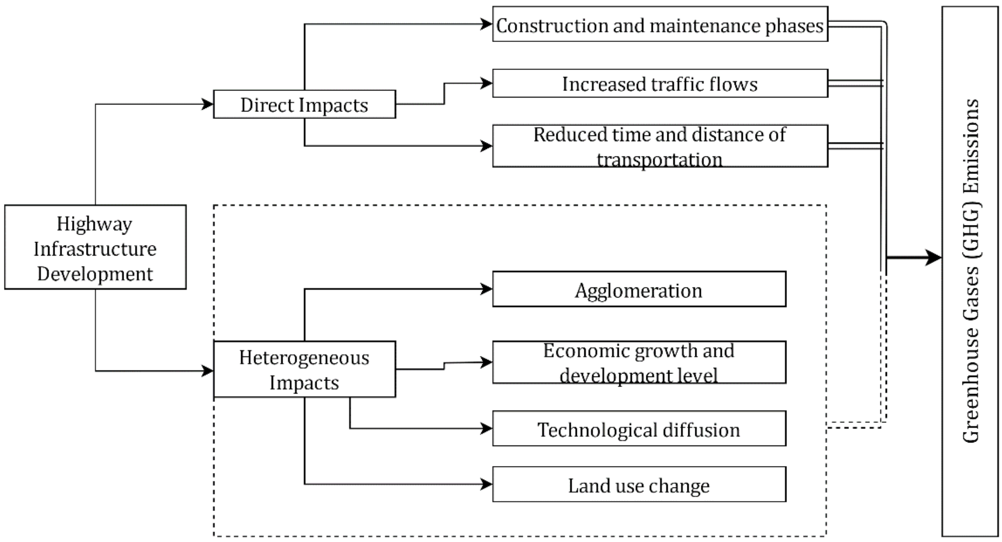
Nonetheless, none of those articles include the environmental costs (benefits) of road investments in the return rate. In other words, the measured return rates are based on the relationship between highway infrastructure investments and economic activity, mainly represented by Gross Domestic Product (GDP) or GDP *per capita*, neglecting any environmental impact from road investments such as increased GHG emissions (Churchill *et al.*, 2021; Xie *et al.*, 2017; Yao *et al.*, 2023), deforestation (Asher *et al.*, 2020), energy efficiency (Lin and Chen, 2020), or ecological footprint (Awad *et al.*, 2023). Disregarding the environmental impacts of highway investments might bias the return rates and directly impact road public policies. Next, we consider the relationship between highway infrastructure development and the environment by focusing on GHG emissions, the most evaluated environmental outcome in transportation studies.

4.2.2. Highway infrastructure and sustainable development

A recent strand of literature has investigated the nexus between road investments and GHG emissions (Emodi *et al.*, 2022; Georgatzi *et al.*, 2020; Ghannouchi *et al.*, 2023; Luo *et al.*, 2018). This relationship is unclear, and there are two opposite views on the effect of highway infrastructure development on GHG emissions (Xu *et al.*, 2022). On the one hand, in the construction and maintenance phases, infrastructure development tends to increase GHG emissions directly by using materials and equipment that tend to be characterized by heavy-duty fuel-intensive equipment and require the use of large quantities of concrete and asphalt (Han *et al.*, 2017; Lin *et al.*, 2017). Once the highway is built, the growth in the road network increases regional accessibility, population mobility, and interregional traffic flows, boosting transportation demand and affecting GHG emissions. On the other hand, some investigations suggest that developing highway infrastructure has a GHG reduction effect by lowering travel time and distance, which decreases GHG emissions during transportation. In addition, transportation infrastructure development might promote agglomeration and technology diffusion, which might support the development of energy savings and emissions reduction.

Following this line of research, some investigations have provided evidence of the road impact heterogeneity on GHG emissions (Churchill *et al.*, 2021; Li and Lu, 2022; Lin *et al.*, 2017; Xiao *et al.*, 2023; Xie *et al.*, 2017; Yao *et al.*, 2023). Figure 4.1 summarizes the mechanisms.

Figure 4.1. The impacts of highway infrastructure on GHG emissions



Source: authors' elaboration.

The most evaluated heterogeneity is related to agglomeration economies. Developed highway infrastructure optimizes the flows of goods and services and the mobility of people within the region, increasing the spatial agglomeration of economic activity through economies of scale and scope. In turn, agglomeration and GHG emissions are strongly correlated. On the one hand, agglomeration tends to increase GHG emissions due to increased production scale and congestion effects. On the other hand, some positive externalities in terms of knowledge spillovers and technological advances might improve energy efficiency and lower energy consumption, decreasing GHG emissions. Studies have found that highway infrastructure expands GHG emissions in the early stages of urbanization and agglomeration. Still, after agglomeration exceeds a threshold, positive externalities from agglomeration are expected to offset the environmentally damaging effects (Xu *et al.*, 2022).

Other researchers have examined different heterogeneity sources, such as economic growth, technological innovation, and tourism, among others (Churchill *et al.*, 2021; Xiao *et al.*, 2023; Xie *et al.*, 2017). Many studies have proven the positive role of highway investments on economic growth (Baum-Snow *et al.*, 2020; Bird and

Straub, 2020; Faber, 2014; Ghani *et al.*, 2014). In turn, economic growth is an important determinant of GHG emissions, as investigations have shown a significant and non-linear relationship between those variables. In addition, transportation infrastructure development fosters the mobility of people, services, and goods, enhancing the spread of knowledge and technology. Technology diffusion impacts GHG emissions and intensity by stimulating human capital formation and higher R&D expenses. Then, the road impacts on GHG emissions are expected to vary according to local level of economic growth and technological innovation.

Besides those investigated moderating variables, other road impact heterogeneities might emerge depending on the local context of GHG emissions. For instance, land use change has been Brazil's most important contributor to GHG emissions. The opening of roads might directly impact GHG emissions by increasing the number of vehicles on the roads and expanding deforestation and illegal land use. Road construction in isolated areas might boost land supply, decreasing land prices and motivating a process of predatory agriculture production wherein landowners have enough incentives to buy new lands instead of improving the existing ones (Carrero *et al.*, 2022; Da Silva *et al.*, 2023; Ferrante *et al.*, 2021; Lima *et al.*, 2022). In addition, the level of deforestation might capture institutional weaknesses related to the environment, which may be translated into a more harmful effect of road infrastructure development on GHG emissions. Then, road investments are expected to present heterogeneous impacts on GHG emissions depending on the level of deforestation and the efficacy of the environmental regulatory framework.

Findings are mixed. Some studies found that road investments increase GHG emissions (Churchill *et al.*, 2021; Emodi *et al.*, 2022; Ghannouchi *et al.*, 2023; Lin *et al.*, 2017; Luo *et al.*, 2018; Xiao *et al.*, 2023; Xie *et al.*, 2017; Yao *et al.*, 2023), while other investigations showed null or negative road impact on carbon emissions (Georgatzi *et al.*, 2020; Ghannouchi *et al.*, 2023; Han *et al.*, 2017; Li and Lu, 2022). In addition, there are heterogenous road impacts on GHG emissions depending on agglomeration, development level and economic growth, and population scale, among other mediating variables.

While this literature has provided important evidence on the relationship between highway infrastructure and GHG emissions, some gaps remain. First, it is hard to interpret how environmentally harmful (or beneficial) road investments are. A way to overcome this issue is calculating a return rate to highway investments considering its effect on GHG emissions, which past studies have not made. Second, most of the investigations have focused on China and European countries, where GHG emissions are mainly generated by the energy and industry sectors. Evaluating the impact of highway investment on the environment in different countries, where GHG emissions depend more on other sectors, such as land use change and agriculture, might be an important contribution to the literature. Third, studies on the nexus between road infrastructure and environmental outcomes at the regional or local levels are scarce, and the existing literature examines the Chinese case.

Evaluating the road impacts on GHG emissions using detailed local-level data might allow the identification of novel heterogeneities in this relationship. This paper seeks to contribute to the literature in those directions.

4.3. Methods

4.3.1. Baseline econometric specification

We intend to evaluate the impacts of highway investment on municipal GHG emissions growth between 2007 and 2018. Our second-stage equation is specified as follows:

$$\Delta Y_{is} = \beta_0 + \alpha * HighwayInvestments_{is} + \beta' X_{is} + \theta_s + u_{is} \quad (1)$$

Where Y_{is} is our dependent variable measured as CO2 equivalent emissions, i represents municipalities, s indicates the states, X_{is} is a vector of control variables, θ_s is a vector of state fixed effects and u_{is} is an idiosyncratic error term. We are interested in α, which measures the impact of highway investment on CO2 emissions. As we take our variables in log form, α is the elasticity of CO2 emissions concerning highway investments.

To estimate the causal impacts of highway investments on CO2 emissions, we adapt the third-step IV identification approach proposed by Medeiros *et al.* (2024). To overcome measurement errors in the road investment variable – due to inefficiencies as corruption, harmful bureaucracy, and poor infrastructure project planning and execution – as well as reverse causality and omitted variable bias – policymakers might target more developed regions wherein the returns to infrastructure investments are higher, or focus on underdeveloped localities to foster regionally balanced economic growth – in the econometric estimates evaluating the road impacts on productivity in Brazilian municipalities, the authors built several instruments related to the propensity of municipalities to receive road interventions. In this paper, the same endogeneity issues may appear whether we have omitted variables affecting environmental outcomes and road placement, which is highly expected (Asher *et al.*, 2020; Emodi *et al.*, 2022; Li and Luo, 2022).

Our preferred specification uses a Non-Random Allocation Index, capturing the propensity of municipalities to receive highway investments as a source of quasi-random variation to road investments. To create the index, Medeiros *et al.* (2024) used the Principal Component Analysis (PCA) method to reduce the data information from three original instruments. The first is the distance from a hypothetical network constructed using the Least Cost Path-Minimum Spanning Tree (LCP-MST) method following Faber (2014). This IV is a global minimization road network connecting the ending and starting points of those roads targeted by the PAC. The rationality behind the LCP-MST instrument is that this hypothetical highway network should affect city outcomes and the spatial allocation of industries only through the actual highway network, conditional on controls. The second original IV follows the Bird and Straub (2020) Brasília experiment approach. The instrument is measured as the distance from targeted central cities to the capital Brasília, and its rationale is that the national Brazilian government in the 1950s and 1960s aimed to connect the whole country having the new capital Brasília as the central point of the network, and municipalities in the way among Brasília and the endpoints were incidentally connected. The third original IV is the distance from the municipality center to the nearest heavy traffic area, which Medeiros *et al.* (2024) named "*potential road intervention areas"* IV. The rationality behind this instrument is that conditional on controls, municipalities already connected by roads in the start period and nearer to "potential road intervention areas" are more likely to (inconsequentially) receive highway investments to reduce traffic levels and accidents in the critical areas and its surroundings. However, conditional on controls, this "luck" at receiving a federal road intervention would be unrelated to economic or political reasons, providing us with a potentially suitable instrument. Finally, we rely on the inconsequential unit approach pioneered by Chandra and Thompson (2000) and exclude likely targeted and central cities. Then, our first-stage regression is specified as follows:

$$HighwayInvestments_{is} = \gamma_0 + \delta * NonRandomAllocationIndex_{is} + \gamma' X_{is} + \varepsilon_{is} \qquad (2)$$

Where $NonRandomAllocationIndex_{is}$ is the instrument. Equations 1 and 2 are estimated using Two Stage Least Squares (2SLS) estimators. Using this econometric approach, we provide evidence of causal highway investment impacts on GHG emissions growth. To guarantee a full comparison with the Medeiros *et al.* (*forthcoming*) economic return rates to highway investments, we also test models including an interaction term between the highway variable and an infrastructure reliance parameter (φ), as follows:

$$\Delta Y_{is} = \beta_0 + \alpha * \varphi * HighwayInvestments_{is} + \beta' X_{is} + \theta_s + u_{is} \qquad (3)$$

If α is positive in Equation 3, municipalities more dependent on road infrastructure are more impacted in terms of GHG emissions growth.

4.3.2. Road heterogeneity econometric specification

A recent strand of literature has provided evidence of the heterogeneous impacts of road investments on environmental outcomes (Churchill *et al.*, 2021; Li and Lu, 2022; Lin *et al.*, 2017; Xiao *et al.*, 2023; Xie *et al.*, 2017; Xu *et al.*, 2022; Yao *et al.*, 2023). This is important as those heterogeneities might bias our baseline estimates and profoundly influence our sustainable return rate to highway investments. Then, we adapt our baseline first and second-stage equations to allow for road impact heterogeneity as follows:

$$\Delta Y_{is} = \beta_0 + \alpha * \varphi * HighwayInvestments_{is} + \lambda' * \varphi * (HighwayInvestments_{is} * Moderators_{is}) + \beta' X_{is} + \theta_s + u_{is} \quad (4)$$

$$HighwayInvestments_{is} = \gamma_0 + \delta * NonRandomAllocationIndex_{is} + \tau' * (NonRandomAllocationIndex_{is} * Moderators_{is}) + \gamma' X_{is} + \varepsilon_{is} \quad (5)$$

Where $Moderators_{is}$ is a vector of moderating variables related to agglomeration economies, technology, deforestation and so forth, which are all included in the vector of control variables as well, and λ' is its respective parameter vector to be estimated. The second stage equation (4) identifies road impact heterogeneity by including an interaction term between the road variable and a moderating variable. To allow identification, we include an interaction term between the instrument and the mediator in the first stage equation (5), wherein τ' represents its parameters vector to be estimated. The other expressions are the same as Equations 1 and 2. From Equations 3 and 4, we can calculate road impact heterogeneity as follows:

$$\frac{\partial Y_{is}}{\partial HighwayInvestments_{is}} = (\alpha * \varphi) + (\lambda * \varphi) * (\overline{Moderators_{is}}) \quad (6)$$

Equation 6 describes the marginal road impact on CO2 emissions. We estimate α and λ directly from Equations 3 and 4. Then, we assume values for $\overline{Moderators_{is}}$ by taking 10%, 25%, median, 75% and 90% sample values for each tested moderator. To estimate the point elasticities, we use tests of nonlinear combinations of parameter estimates following the "delta method" (Fieveson, 1999).

4.3.3. Data

4.3.3.1. GHG Emissions

Our primary dependent variable is CO2 equivalent emissions (in tons), which we extract from the System for Estimating Greenhouse Gas Emissions (SEEG). All gases were converted to CO2 equivalent GWP-AR5. The SEEG platform is a 46-year-long dataset of greenhouse gas emissions (GHG) in Brazil (1970–2015), providing more than 2 million data records for the Agriculture, Energy, Industry, Waste, and Land Use Change Sectors at national and subnational levels. The SEEG dataset was developed by the Climate Observatory, a Brazilian civil society initiative, based on the Intergovernmental Panel on Climate Change (IPCC) guidelines and the Brazilian National Inventories embedded with country-specific emission factors and processes, raw data from multiple official and non-official sources, and organized together with social and economic indicators. Due to the SEEG's highly disaggregated information, we can stratify municipal GHG emissions into road, energy, land use change, and agriculture sectors and use them as additional dependent variables. A detailed description of the SEEG platform and methodologies can be found in Azevedo *et al.* (2018).

4.3.3.2. Highway infrastructure measures and instruments

Our interest variable is the sum of federal highway investments between 2007 and 2018 by municipality. We get highway investment data from Medeiros *et al.* (2024). The authors created a municipal-level federal road investment dataset by combining the PAC highway investment data with the georeferenced National Highway System (SNV) from the National Highway Infrastructure Department (DNIT). We also try two additional road variables as robustness checks to measure measurement errors in our road investment measure. The first one is a dummy variable assuming value one if the municipality received a road investment during the PAC period and zero otherwise. The second one is the road length growth rate between 2006 and 2018. In this case, we use 2006 data from the 2007 National Transport Logistics Plan (PNLT) and 2018 data from DNIT[22]. To maintain comparability with the economic return rate calculated by Medeiros *et al.* (*forthcoming*), we use their infrastructure reliance parameter (φ), measured as the share of the municipal intermediate consumption related to the land transportation sector. The φ data sources are the Annual Social Information Report (RAIS/Ministry of Labor) and the 2010 National Input-Output (I-0)(IBGE).

In addition, we also rely on Medeiros *et al.* (2024) as the source of our instrumental variables. We get the Non-Random Allocation Index as our main IV and their three original instruments as robustness checks. In addition, we also get some cost-related IVs (Cost Index 1 and 2) related to environmental, geographic, and human physical infrastructure project costs to run additional tests. The indexes were also created by using the PCA technique, reducing data information from original variables as the share of hilly areas in the total area, the share of urban infrastructure building in the total area, the share of legally protected environmental areas in the total area and the application of environmental embargoes.

4.3.3.3. Moderating variables

We include an extensive set of controls to avoid omitted variables bias following the specialized literature on road infrastructure and GHG emissions (Churchill *et al.*, 2021; Emodi *et al.*, 2022; Georgatzi *et al.*, 2020; Ghannouchi *et al.*, 2023; Han *et al.*, 2017; Li and Lu, 2022; Lin *et al.*, 2017; Luo *et al.*, 2018; Sharif and Tauqir, 2021; Xiao *et al.*, 2023; Xie *et al.*, 2017; Xu *et al.*, 2022; Yao *et al.*, 2023) and adapting for Brazilian features. First, we describe some variables that will be used as both controls and moderators. We include population to control and moderate for city scale and agglomeration effects. We also try population density and the share of

[22] We can likely observe measurement error in the road length variable as well, as the PNLT and DNIT files are not fully comparable. In addition, there is methodological variations over the years in relation to road classifications as federal, state level and so forth. Then, this variable should be used with caution.

road sector CO_2 emissions in the total CO_2 emissions as robustness checks. Second, we include GDP *per capita* to control for the municipal development level. Third, we include the ratio between residential capital and occupied population as a *proxy* for technological innovation. Finally, we include deforestation variation between 1996 and 2006 as a control for the municipal propensity to raise land use change GHG emissions, the country's main source of GHG emissions.

4.3.3.4. Additional controls

As additional controls, we include the initial (2007) level of GHG emissions to control for level and convergence effects. We also include GDP per capita squared to control for a potential environmental Kuznets Curve. We include the share of the municipality exports in the national exports as a control for trade specialization. Gini Index controls for income inequality. Institutional Quality is inserted using the Index of Municipal Institutional Quality (IQIM). Human capital is included as the share of workers with graduate education. We control for complementary and substitute infrastructure by including the Euclidean distance from the municipality center to the nearest state road, port, and railroad. To guarantee the suitability of our instruments, we also include the distance to Brasilia and the number of railway stations in 1920 as controls, as Medeiros *et al.* (2024) relied on historical data to construct some of their IVs. A brief description of the variables used and their sources can be found in Table A1, and descriptive statistics can be seen in Table A2 in Appendix A.

4.4. Econometric results and discussion

4.4.1. Baseline estimates

Table 4.1 presents our baseline econometric results by estimating Equations 1 and 2[23]. In the first five columns, we use our highway investment measure as the interest variable. In columns 6-10, we multiply our road variable by the road infrastructure reliance parameter (φ) following Fernald (1999). We estimate the road investments' impact on CO_2 emissions considering the full sample (columns “All”) and the road, energy, land use change, and agriculture sectors separately. The Non-Random Allocation Index strongly predicts road investments and is a pretty suitable IV, as indicated by the high F Statistic values. Regarding the second stage regressions, we find a positive relationship between road investments and GHG emissions for the full sample as well as for the road, energy, and land use sectors. We found no significant road effects on agriculture GHG emissions.

[23] Tables B1 and B2 in Appendix B presents OLS regressions results as comparison estimates.

Table 4.1. Federal Highway Investments and CO2 Emissions Growth (2007-2018): 2SLS IV Regressions

	1	2	3	4	5	6	7	8	9	10
Second stage	All	Roads	Energy	Land Use	Agriculture	All	Roads	Energy	Land Use	Agriculture
Log Highways Investments	0.0249***	0.1335***	0.1157***	0.0532***	-0.0030					
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)					
Log Highways Investments * φ						0.5770***	3.0916***	2.6802***	1.2362***	-0.0702
						(0.18)	(0.51)	(0.49)	(0.36)	(0.14)
First stage										
Non-Random Allocation Index	-0.4702***	-0.4808***	-0.4826***	-0.4759***	-0.4797***	-0.0203***	-0.0208***	-0.0208***	-0.0205***	-0.0207***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	5142	5142	5142	5142	5142	5142	5142	5142	5142	5142
KP Wald F Statistic	349.317	359.100	360.198	356.129	360.001	332.469	343.528	344.349	338.341	340.560
R^2	0.23	0.51	0.52	0.22	0.15	0.23	0.50	0.52	0.21	0.15

All regressions include the following set of control variables: CO2 emissions in 2007; state fixed effects; population; GDP *per capita*; GDP *per capita*, square; capital-labor ratio; exports share; 1996-2006 deforestation; Gini index; institutional quality; human capital; distance to the nearest state road; distance to the nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília. Robust standard errors are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

We can interpret our findings in terms of elasticity. More directly, a 1% increase in highway investments increases CO_2 emissions by 0.025%. As expected, the elasticity is larger for the road and energy sectors, suggesting that a 1% increase in road investments raises road and energy CO_2 emissions by 0.134% and 0.116%, respectively. These results corroborate several studies that found a damaging effect of highway construction and improvement on the environment, especially in the urban and road-related context (Churchill *et al.*, 2021; Ghannouchi *et al.*, 2023; Lin *et al.*, 2017; Luo *et al.*, 2018; Xie *et al.*, 2017; Yao *et al.*, 2023).

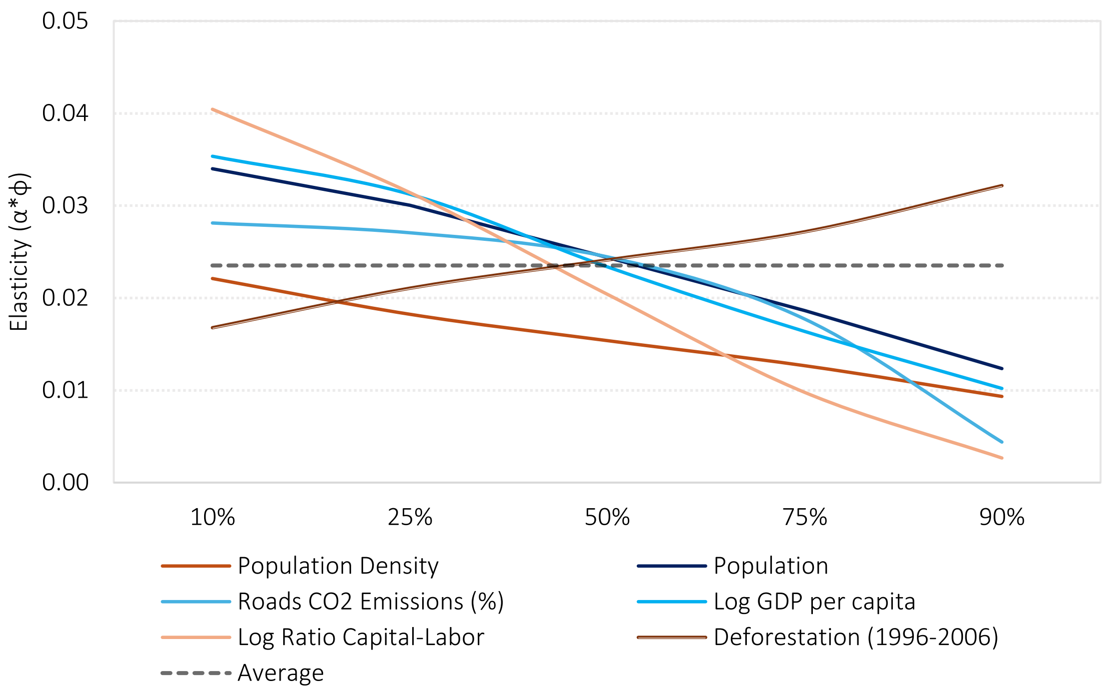
Moreover, results point out a positive and significant indirect road effect on land use change CO_2 emissions. These novel findings might be explained in some ways. Opening new highways in isolated and previously environmentally protected areas might expand land supply. Therefore, land prices drop because of the expanded land offer, and landowners might be more prone to buy new lands instead of investing in improving the productive efficiency of the existing ones. This might lead to a process of land abandonment, predatory agriculture and illegal extractivism, with consequent deforestation and destruction of fauna and flora (Carrero *et al.*, 2022; Da Silva *et al.*, 2023; Ferrante *et al*, 2021; Lima *et al.*, 2022). As a result, we might expect an increase in CO_2 emissions related to land use change from road investments.

4.4.2. Road impact heterogeneity

In this section, we evaluate the potential heterogeneous effects of road investments on GHG emissions. To do this, we interact our road variable with some interesting moderating variables following Equations 4 and 5. Then, we calculate point GHG emissions elasticities concerning highway investments by applying Equation 6. Figure 4.2 exhibits the results. Full estimation results can be seen in Table C1 in Appendix C.

The first moderators we analyze are related to agglomeration economies and population scale effects (Lin *et al.*, 2017; Xiao *et al.*, 2023; Xie *et al.*, 2017; Xu *et al.*, 2022; Yao *et al.*, 2023). We test interactions between our road variable and population, using population density as a robustness check. As an additional test, we use the share of the road sector's CO_2 emissions in relation to the total CO_2 emissions to represent the importance of the road sector in the municipality economy and emissions, as well as to identify places wherein high traffic congestion is expected. Findings point out that the positive impact of highway infrastructure improvement on CO_2 emissions is higher for lower levels of our moderating variables. For instance, a 1% increase in road investments raises CO_2 emissions by 0.034% in the bottom 10% of the population, while the elasticity is 0.012% in the upper 10% of the same variable. The same rationality holds for population density and the share of road sector CO_2 emissions. These results align with investigations reporting a significant moderating effect of agglomeration and population scale on the relationship between transportation development and carbon emissions. Agglomeration, as the most direct manifestation of the positive externality of highway infrastructure, is the core driver of rapid regional economic growth and supports the development of energy savings and emission reductions in society. As municipalities reach a certain level of urbanization and agglomeration, the effects of roads on CO2 emissions become less harmful (Lin *et al.*, 2017; Xie *et al.*, 2017; Xu *et al.*, 2022).

Figure 4.2. Federal Highway Investments and CO2 Emissions Growth (2007-2018) - Elasticity (α*φ): Heterogeneous Impacts



Source: authors' elaboration.

Next, we interact the road variable with GDP *per capita* and the capital-labor ratio to capture heterogeneities regarding local development levels and technology innovation, respectively (Churchill *et al.*, 2021; Xie *et al.*, 2017). Like the agglomeration economies and population scale moderators, the road impact on CO2 emissions increases with the levels of GDP *per capita* and technology. The variation is more pronounced in the technology mediator, suggesting an α equal to 0.04% in the bottom 10%, 0.01% in the upper 25%, and a non-significant (nearly zero) effect in the upper 10%. These findings indicate a greater polluting impact of roads in less developed locations, probably due to the construction of new roads and the expansion of new markets. As Medeiros *et al.* (*forthcoming*) found, the effects of highways on the local economy tend to be greater in poorer locations. Technology is positively correlated with economic development. In this sense, we expect roads to expand CO2 emissions through economic growth and technological innovation channels in the initial stages of development.

Next, we include an interaction term between the road variable and the deforestation variation in the past ten years to the PAC. Results show that the damaging highway investments effects on the environment enlarge while deforestation in the recent past increases. A higher level of deforestation might represent national and local institutional weaknesses, allowing the purchase of new lands at lower prices and its illegal use - such as land grabbing. By opening new roads, landowners may access new lands that were not available before, turning to a process of predatory agriculture production and widespread deforestation (Carrero *et al.*, 2022; Da Silva *et al.*, 2023; Ferrante *et al.*, 2021; Lima *et al.*, 2022). This finding puts some caution on the role of road policies on sustainable development in Brazil, especially in the Brazilian Amazon municipalities, as the region has suffered from massive deforestation in the last decades.

4.4.3. Robustness checks

In this section, we present some robustness checks to increase confidence in our main results described so far. First, we used a highway investment flow measure as our preferred variable. However, several studies advocate against measuring infrastructure in monetary terms as inefficiencies in project planning and design, as well as corruption and flawed bureaucracy, might turn investments ineffective in terms of building and implementing infrastructure, especially in developing economies. In other words, monetary variables might not represent effective infrastructure appropriately (Calderón and Servén, 2014; Kenny, 2009; Straub, 2011). This issue is alleviated as we used an IV identification approach dealing with endogeneity, but some bias may remain. As robustness checks, we use a dummy variable assuming value one if a municipality received a PAC highway intervention and zero otherwise. In addition, we try road length growth between 2007 and 2018 as an interest variable following a vast strand of literature[24] (Baum-Snow *et al.*, 2020; Duranton *et al.*, 2014; Foster *et al.*, 2023a, 2023b; Straub, 2011). Results can be seen in Table D1 in Appendix D. Findings corroborate our baseline estimates, suggesting a positive impact of road infrastructure on CO_2 emissions. In addition, we find α equal to 0.10% using the road length variable, which is quite in line with the elasticity of 0.08% estimated by Xie *et al.* (2017).

Second, we try additional IV combinations to validate our identification strategy. Table D2 in Appendix D presents the results. In Column 1, we include two infrastructure cost indexes following the three-step IV identification approach by Medeiros *et al.* (2024). In Columns 2-4, we use the three original non-random allocation instruments instead of the Non-Random Allocation Index. In Columns 5-7, we include the two cost indexes jointly with the original non-random placement IVs. Results remain almost unchanged, presenting a stable elasticity.

[24] It is important to mention that this measure likely has measurement errors (perhaps more problematic than the monetary measure) due to changes in methodology and issues related to spatial disaggregation. Therefore, the results should be taken with caution.

Third, we run the same baseline models considering $CO2$ emissions in 2018 levels instead of growth rates as dependent variables. Results (Table D3 in Appendix D) are preserved. Next, we try additional robustness checks to alleviate concerns about CO2 emissions regional heterogeneity. First, we exclude all municipalities belonging to Amazon states. Those localities have suffered the most from deforestation in the past decades, and very high amounts of CO2 emissions related to land use change might affect our baseline estimates. Second, we drop municipalities in the state of Pará. Pará received emblematic road buildings in environmental terms, some crossing extensive native people lands and generating huge environmental damages and land conflicts (Medeiros *et al.*, 2024). Third, we exclude municipalities of the state of São Paulo to alleviate issues related to high urbanization and development levels, which might impact our estimates due to a substantial share of CO2 emissions related to the road and energy sectors. Finally, we estimate the road effects on CO2 emissions growth pertaining to the land use change sector by excluding municipalities of the Amazon states. If our baseline estimates capture a regional road effect in the Brazilian Amazon, this robustness check should not present a significant parameter to road investments. Results can be seen in Table D4 in Appendix D. Findings remain, corroborating our baseline estimates. In unreported estimates, we also try the limited information maximum likelihood (LIML) and the generalized method of moments (GMM) estimators, and the results are unchanged.

4.5. Including Sustainability into the Return Rate to Highway Investments

4.5.1. The CO2 Emissions Return (Discount) Rate to Highway Investments (ERR) and the Sustainable (and Equitable) Return Rate to Highway Investments (SRR and SERR)

In this section, we provide a novel measure we call Sustainable Return Rate to Highway Investments (SRR). To do this, we take the (economic) Return Rates (RR) calculated by Medeiros *et al.* (*forthcoming*) – which consider the road impact on productivity measured as GDP *per capita*, i.e., economic returns[25] – and discount from it our CO2 emissions Return Rate (ERR). To calculate the ERR, we adapt the return rate formula used by several studies (Fernald, 1999; Medeiros *et al.,* 2021; Medeiros *et al.,* 2024*, forthcoming;* Wang *et al.*, 2020) as follows:

$$ERR_r = \alpha * \varphi_r * \frac{CO2Emissions_r^{SCC}}{HighwayStock_r} \quad (7)$$

[25] The formula used by Medeiros *et al.* (2024) to calculate the RR is: $RR_r = \alpha * GDP_r / RoadStock_r$, where α is the road elasticity in relation to GDP *per capita*, which multiplies the ratio between regional GDP and the road stock.

Where $CO2Emissions_r^{SCC}$ is the total CO2 emissions in monetary terms, $HighwayStock_r$ is the stock of roads in monetary terms and r represents Brazilian Immediate Geographical Regions (RGI).

We follow Medeiros *et al.* (*forthcoming*) in some steps to construct our ERR and to guarantee comparability with their RR. First, we alleviate issues with outliers by taking the decile average values of the ratio between CO2 emissions and the highway stock and then applying these averages to each municipality. Additionally, we exclude municipalities in the top and bottom 1% when calculating those averages to reduce measurement error bias from extraordinarily high and low ratio values. Second, we include the infrastructure reliance parameter (φ_{is}) to allow local road dependence heterogeneity to work. Third, we aggregate the municipality values at the RGI level by taking the average values of φ and the ratio between CO2 emissions and road stock, i.e., dividing those variables by the number of municipalities in each RGI. This third step is essential to policy implications as we do not expect the Brazilian Federal Government to target specific municipalities in allocating roads. The 510 RGIs are groups of municipalities in the urban network sharing a common local urban center as their basis, being constructed by the IBGE. Its design considers the connection of nearby cities through relationships of dependency and the movement of the population in search of goods, services, and employment opportunities. Then, the RGIs are closely related to transportation goals and can be seen as a reasonable spatial scale regarding national highway public policies.

To construct the road stock variable, we follow Medeiros *et al.* (2021a) and Medeiros *et al.* (2024*, fortchoming*b) by using the Frischtak and Mourão (2017) sectoral estimates for the Brazilian road stock. The authors found a road stock of around R$ 594 billion in 2023 values. Next, we use georeferenced road data from the 2007 National Transport Logistics Plan (PNLT) to calculate the road length by municipality. We multiply single lanes by one and duplicated lanes by 2 to control for road quality and scale in our stock measure. Then, we divide the total road stock in monetary terms by our physical measure of road length to generate the monetary value by kilometer of road. Finally, we multiply this value by the road length of each municipality, which gives us our local road stock variable.

To generate our ERR, we also need to quantify CO2 emissions in monetary terms. To this end, we use the measure of the Social Cost of Carbon (SCC). The SCC is an estimate of the cost, in dollars, of the damage done by each additional ton of carbon emission. SCC estimates mostly evaluate the impacts of carbon emissions on health outcomes, agricultural production, and property values.

However, there is no consensus on the SCC value to be applied. Then, we use some benchmark SCCs to ensure consistency in our results. The first SCC we use is the Brazilian Government one (Ministry of Economy, 2022). The Brazilian government's SCC of around US$ 31 was mainly guided by a literature review considering several studies estimating the SCC worldwide (Nordhaus, 2016). The Brazilian Government's SCC is in line with the Ricke *et al.* (2018) median SCC for Brazil of around U$$ 24.2 – the authors calculated country-level SCC values for several developed and developing economies –which we also use as SCC in our ERR calculation. Recent studies have established substantially larger SCC values considering different contexts and methodologies (Rennert *et al.*, 2022). For instance, the US Government SCC – one of the most relevant SCCs guiding carbon pricing and environmental policies around the globe – is around US$ 51. Even so, several academics consider the North American SCC low, suggesting values above US$ 100. In this sense, we also consider the SCC of U$$ 113 proposed by the United Nations Environment Programme (UNEP, 2014), also identified in the UK Government's Stern report as the central, business-as-usual scenario value. Finally, we convert the SCCs to the Brazilian currency (R$) using an exchange rate of R$/US$ 5.17.

Importantly, we have demonstrated some significant heterogeneous road impacts on CO2 emissions. Whether Brazilian municipalities and regions present high variability in the values of the moderator variables, we can expect some bias in our ERR by taking a single average α value. To alleviate this issue, we adapt Equation 7 using the road heterogeneity impact results as follows:

$$ERR_r = ((\alpha * \varphi_r) + (\lambda * \varphi_r) * Moderator_r)) * \frac{CO2Emissions_r^{SCC}}{HighwayStock_r} \tag{8}$$

Where λ is the interaction term parameter allowing the road impact heterogeneities to exist, and $Moderator_{is}$ is the moderator values taken by the municipal average by RGI.

Finally, we calculate our Sustainable Return Rate to Highway Investments (SRR) as follows:

$$SRR_r = RR_r - ERR_r \tag{9}$$

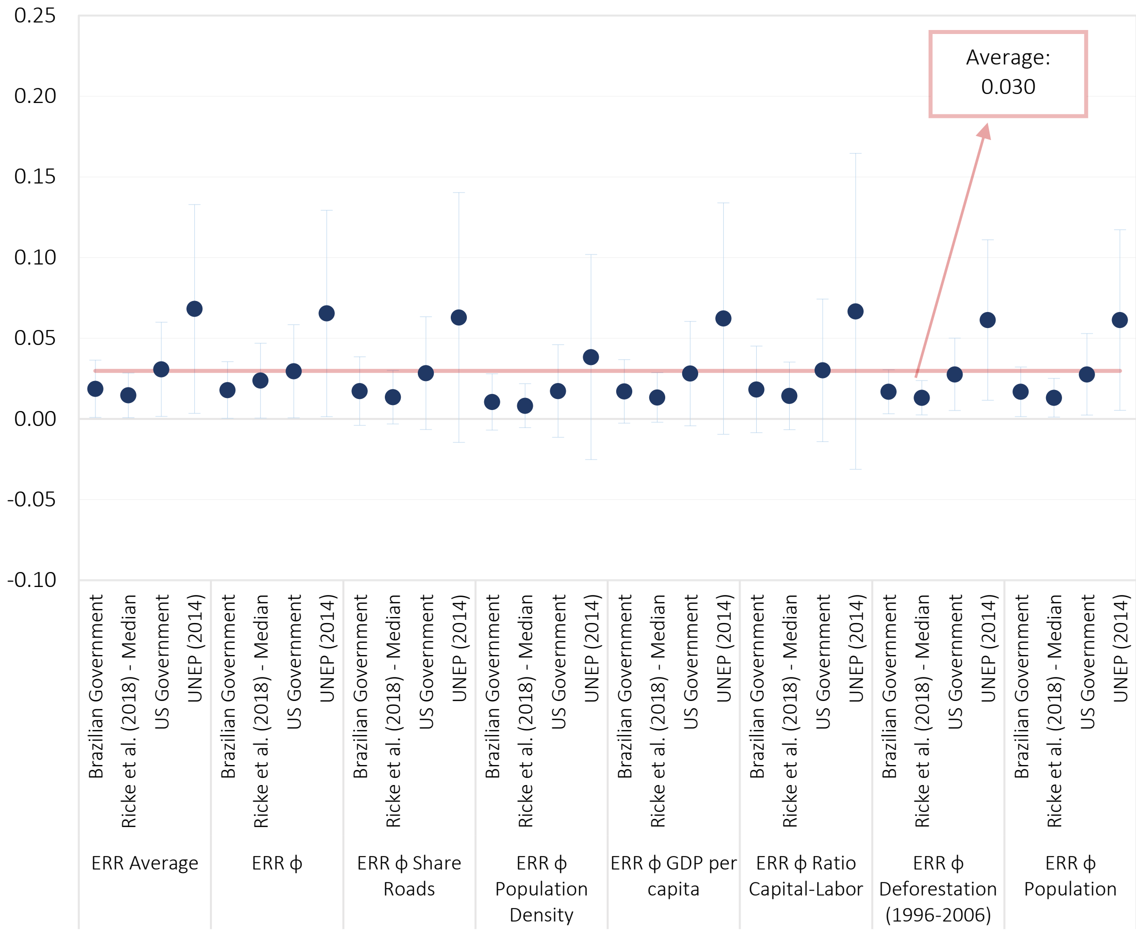
Equation 9 shows two opposite sides of road policies. In other words, the higher the ERR, the lower the positive economic returns of road investments to society.

4.5.2. Results and policy implications

Figure 4.3 shows the ERR results. We calculate several ERRs by trying different SCC values and varying our parameters, following the results in Sections 4.1 and 4.2. Our ERR ranges from 0.01 – using the Rick *et al.* (2018) SCC and the parameters following the population density moderator specification – to 0.07 – taking the UNEP (2014) SCC and the parameters from the average α specification. To establish a benchmark for the ERR, we suggest taking the average value of all ERRs exhibited in Figure 4.3, indicating an average ERR of 0.03 (3.0%).

The average economic return rate (RR) by Medeiros *et al.* (2024, *forthcoming*) is around 20%. Discounting our ERR from the average RR implies an SRR ranging from 13% to 19% in Brazil. On average, we find a high SRR of 17%. This result corroborates the consensus on the deep precarity of the Brazilian transportation infrastructure sector, even considering environmental damages.

Figure 4.3. CO2 Emissions Return Rate to Highway Investments (ERR) under different Social Costs of Carbon (SCC) and road impact heterogeneities

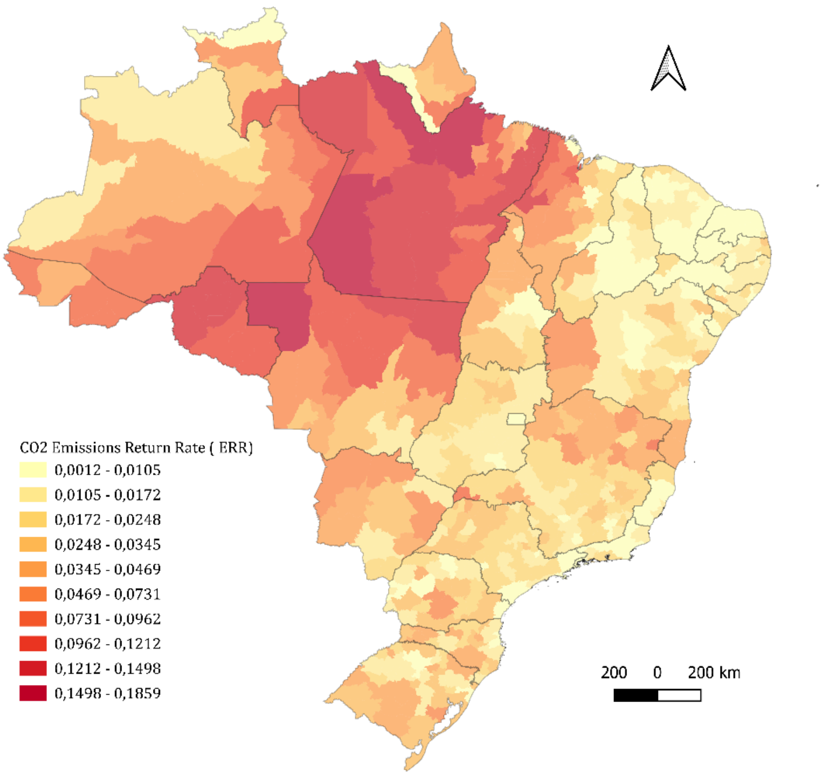


Source: authors' elaboration.

Nonetheless, Brazil presents huge regional heterogeneities in terms of CO2 emissions and road dependence and looking at those features might reveal some spatial inequalities in the ERR and SRR. Figure 4.4 shows the ERR at the regional scale. We can observe a substantial number of regions presenting low ERR values between approximately zero and 0.025. However, for an important part of RGIs in the north and part of the Mid-West regions – more specifically, in the Brazilian Amazon area –, our results indicate ERRs above 0.07, reaching peak values of around 0.19. Highway investments might constitute an environmentally damaging policy tool for those with high ERR values.

To better elucidate how the environmental and economic issues of highway policies are operating, we display the SRR in Figure 5. In Figure 4.5 (a), we show the Sustainable Return Rate to Highway Investments (SRR) considering the average RR calculated by Medeiros *et al.* (*forthcoming).* In the RR Average, the authors consider the road impact on productivity equal to all units. In Figure 4.5 (b), we show the Sustainable and Equitable Return Rate to Highway Investments (SERR), which considers the RR Efficient & Road Specialized & Redistributive & Equative measured by Medeiros *et al.* (*forthcoming*). In this second return rate, the authors allowed the road impact on productivity to vary by units and found that the road investment profitability is higher for less developed and poorer infrastructure-endowed places. Then, we evaluate the road return in terms of economic profitability (RR), weighting by social conditions considering equity features (RR Efficient & Road Specialized & Redistributive & Equative), and sustainability (ERR). The SERR is our preferred estimate as it deals with a broader range of road policy characteristics, going beyond the widely evaluated economic issue.

Figure 4.4. CO2 Emissions Return Rate to Highway Investments (ERR): Brazilian RGIs

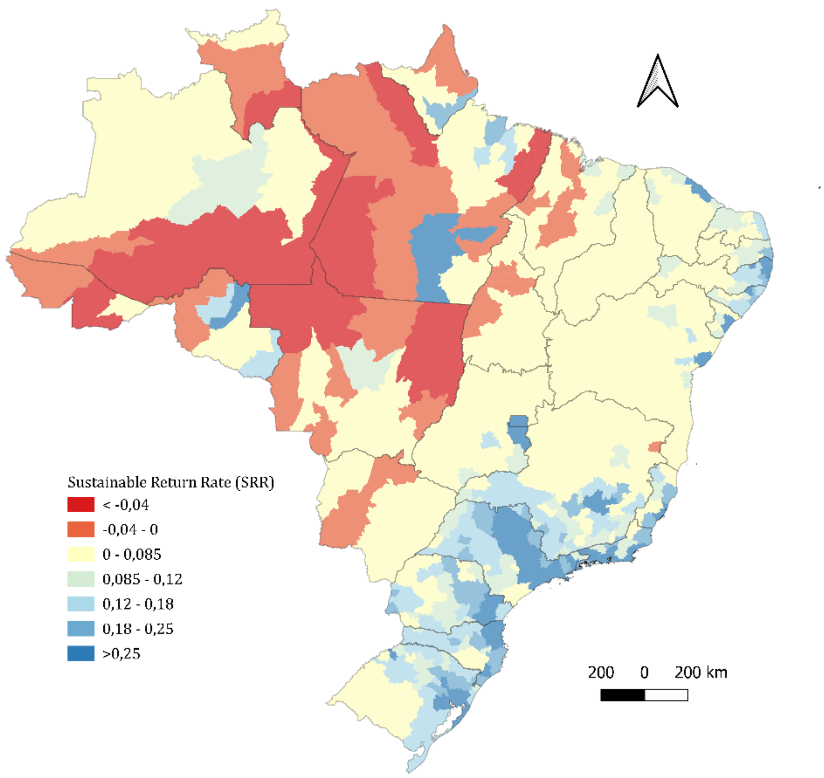


Source: authors' elaboration.

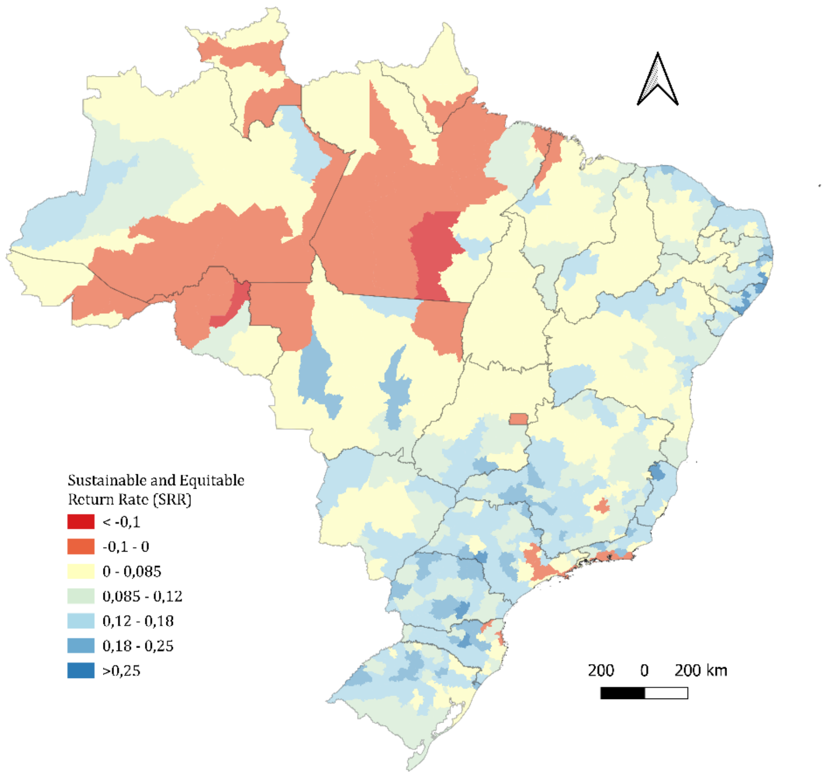
While we observe a high average SRR, Figure 4.5 shows some critical regional disparities in Brazil. First, we can observe positive SRRs for most of the country, as expected due to historical bottlenecks in the Brazilian road sector. However, many RGIs in the north and Mid-West regions present negative SRRs, implying that the environmental costs are higher than the economic benefits of constructing and improving roads in those localities.

Figure 4.5. Sustainable Return Rates to Highway Investments: SRR (a) and SERR (b)

(a)



(b)



Source: authors' elaboration.

When considering the SERR, the number of non-profitable RGIs drops as the economic (and equitable) return is higher for the poorer places, especially in the North and Northeast regions. Evaluating the SERR, we have a larger number of RGIs presenting return rates above 8.5%, the cut-off rate following the Social Discount Rate (TSD) calculated by the Brazilian Ministry of Economy (2021). Even so, some unprofitable and environmentally vulnerable RGIs remain.

Finally, we evaluate the sensitivity of our ERR and SERR measures to land use change GHG emissions, Brazil's leading source of GHG emissions. To this end, we provide some naïve counterfactual exercises supposing drops of 25%, 50%, 75%, and 100% in land use change GHG emissions and recalculate our ERR and SERR. Results are described in Table E1 in Appendix E. Our ERR decreases from the average of 3.0% to 2.54% and 1.42%, supposing a 25% and 100% reduction in land use GHG emissions, respectively. Consequently, our SERR increases from the average of 17% to 17.46% and 18.58%, taking the 25% and 100% reduction in land use GHG emissions, respectively.

Additionally, we generate a new SERR considering the energy sector CO2 emissions[26] (Figure E1 in Appendix E). This exercise aims to avoid the damaging road impacts from deforestation and agriculture, which might be somewhat out of control of the transport sector authorities such as the Ministry of Transport and the DNIT. Then, we restrict the emissions more directly related to the highway improvements, as those are strictly associated with increased traffic flows and urban activity. In this case, we consider the elasticity of the energy sector CO2 emissions growth concerning highway investments to be equal to 0.12, as shown in Table 4.1. The average ERR under the energy sector CO2 emissions analysis is around 1.3% (less than half of the ERR considering emissions from land use change and agriculture), while the average SERR is close to 18.7%. It is important to mention that several RGIs become economically and environmentally profitable when we evaluate only the energy sector's CO2 emissions. This result indicates complementary policies' critical role in preventing deforestation and preserving and restoring the environment, especially in the Amazon region.

Our findings regarding sustainable return rates to highway investments have important policy implications. First, the average return rate to road investments is high even considering the environmental issue, indicating a widespread need to develop the Brazilian transportation sector. To reduce our average SRR of 17% to the threshold of 8.5%, Brazil would need two times more highways, which implies a road stock of 14% of national GDP, in line with Frischtak and Mourão (2017) and Medeiros *et al.* (2021). Second, the environmental damage from roads is more pronounced in less populated and poorer localities, which coincides with some critical areas in the Brazilian Amazon. For some of those RGIs, we can observe negative SRRs and SERRs, suggesting that the economic benefits are not offsetting

[26] We calculate the energy sector ERR following results in Columns 3 and 8 in Table 1. Then, we use the ERR Average and ERR φ specification as in Figure 3.

the raising in environmental costs from road-related GHG emissions. Then, public road policies must be implemented jointly with environmental tools to ensure environmental preservation and recovery. Third, even if moderate, we found a positive average road impact on CO2 emissions, implying a discount on the economic return rate to road investments.

Additional public policies might be essential to alleviate those harmful road impacts. For instance, taxes and subsidies for clean technologies such as electric vehicles and energy systems might make them more attractive. Once those technologies achieve a certain level of production scale, costs tend to fall, and the incentives to produce and use clean technologies become high enough (Greene *et al.*, 2014; Santos, 2017). The same might hold for research and development (R&D) expenses in clean technology. Finally, improving the institutional and regulatory environmental framework is critical, especially for road project design, execution, and evaluation in environmentally vulnerable areas. Improving project governance and coordinating transportation and environmental institutions are vital issues.

4.6. Concluding remarks

We evaluated the impact of highway investments on GHG emissions in Brazilian municipalities during the PAC period (2007-2018). Using an IV identification strategy dealing with the non-random allocation of roads, we find an increasing effect of roads on CO2 emissions, showing that a 1% rise in road investments expands CO2 emissions by 0.025%. This damaging effect of road investments on the environment holds for the road, energy, and land use change sectors. We also found important heterogeneous road impacts on CO2 emissions depending on agglomeration, population scale, deforestation, and technology. In short, less agglomerated and populated and poorer localities are more adversely affected by road investments. We detected a new transmission channel from road investment to CO2 emissions from deforestation, proving that municipalities with higher deforestation in the previous period to the PAC suffered more from the damaging effects of highways on the environment. Findings are robust to different specifications, varying dependent and independent variables and instruments, excluding groups of municipalities, and changing estimators.

From this, we calculated an average CO2 Emissions Return Rate to Highway Investments (ERR) of 3.0%, implying a discount on the economic benefits of road investments proved in past studies. Next, we measured a Sustainable Return Rate to Highway Investments (SRR) of around 17%, indicating a widespread need to develop the Brazilian transportation sector. It is essential to note the existence of deep regional heterogeneities in Brazil, wherein we can observe negative SRRs and SERRs for some regions – especially in the Brazilian Amazon–suggesting that the economic benefits are not offsetting the raising in environmental costs from road-related CO2 emissions in those places.

While we contribute to the empirical literature on infrastructure and development in several ways, some gaps remain. First, we evaluated just one outcome in a wide range of environmental factors potentially impacted by roads. Future research might expand our study by focusing on deforestation, energy efficiency, water pollution, and ecological footprint, among others. Second, a more detailed analysis of the moderating role of environment-related institutions on the nexus between highway investments and GHG emissions might provide important and novel evidence to the literature, especially in countries wherein land use change and agriculture are relevant contributors to GHG emissions. Third, differentiating the short-run and the long-run environmental impacts of road investments may provide important policy implications in terms of pollution from material and equipment in the construction phase versus the environmental damage caused by the increased traffic flows when the highway is already built.

5. BRINGING HIGHWAY INVESTMENTS MORE EFFICIENT, INCLUSIVE, AND SUSTAINABLE: establishing priorities for the development and evaluation of regionalized road policies in Brazil

Abstract

Infrastructure investments are crucial for economic growth, social equity, and environmental sustainability. However, in many cases, such investments prioritize economic returns over social and environmental consequences, leading to regional disparities and environmental degradation. This study proposes an empirical strategy to classify priority regions for highway investments in Brazil, considering economic, social, and environmental issues arising from road investments. We apply this proposal to different spatial scales, allowing us to think about local, regional, and national road policies. Our findings offer novel inputs for policymakers, technicians, financial institutions, and civil society in shaping efficient, equative, and environmentally conscious road policies. We then conduct an *ex-post* evaluation of the Growth Acceleration Program (PAC)(2007-2018) to demonstrate how our framework can be applied for public policy purposes. From this, we find that the program generated meaningful economic returns, but this return could have been 30% higher if the government had focused on our priority regions. New road policies may achieve better results by directing investments to *win-win* regions with the potential to increase economic growth, reduce inequalities and mitigate environmental damage.

Keywords: transportation investments; efficiency; redistribution; equity; sustainability.

5.1. Introduction

Infrastructure investment is a powerful policy tool for fostering economic growth by increasing productivity, expanding internal and external markets, and intensifying agglomeration economies (Aschauer, 1989; Asher and Novosad, 2020; Bird and Straub, 2020; Coşar *et al.*, 2022; Roberts *et al.*, 2020). These investments might also bring social benefits by alleviating poverty and inequalities within and between regions (Medeiros and Ribeiro, 2020; Medeiros *et al.*, 2021a, 2022). In addition, infrastructure interventions affect several environmental outcomes, such as greenhouse gas (GHG) emissions, deforestation, energy efficiency and savings, the ecological footprint, and so forth (Awad *et al.*, 2023; Churchill *et al.*, 2021; Emodi *et al.*, 2022; Lin and Chen, 2020; Xu *et al.*, 2022). In this sense, infrastructure policies might achieve *win-win* combinations distinguished by promoting inclusive and sustainable economic growth.

Nonetheless, these best arrangement scenarios in economic, social, and environmental terms are unlikely to occur. In most cases, infrastructure investments are driven by expected economic returns, overlooking socio-environmental road components (Alam *et al.*, 2022; Calderón and Serven, 2014; Laird and Venables, 2017; Quadros and Naci, 2015; Straub, 2011; Welde and Tveter, 2022). Wealthier regions will likely receive greater private resources because of their higher predictable profitability. On the other hand, when infrastructure investments are lacking and the public budget for capital expenditures is constrained – as it occurred in Brazil in the past decade–, underdeveloped and geographically more inaccessible regions might be left behind due to their smaller economic returns and higher environmental costs. In this context, setting priorities is critical in maximizing economic returns, reducing regional inequalities, and mitigating environmental damage from road development, especially in developing country scenarios marked by scarce infrastructure investment levels.

In this section, we propose a novel empirical approach to classify priority regions for highway investments in Brazil. Unlike previous studies focusing on economic concerns (Fernald, 1999; Li *et al.*, 2017; Medeiros *et al.*, 2021b; Wang *et al.*,2020), we target places where economic, social, and environmental issues work together. The main aim of this chapter is to translate our empirical findings into a policy-making-oriented approach, providing easy-to-interpret inputs for policymakers in planning, designing, financing, and evaluating efficient, redistributive, and sustainable public road policies across the country. Specifically, we offer a regionalized framework containing unique inputs aimed at improving the effectiveness of national highway investments in light of the Growth Acceleration Program (PAC).

To do this, we reassess the findings obtained in Chapters 2, 3, and 4, estimating the Sustainable and Equitable Return Rates to Highway Investments (SERR). In Chapter 2, we built a novel econometric identification strategy to estimate the causal impacts of road investments on economic activity. Using this econometric approach, we calculated some efficient, road-specialized, redistributive and equative return rates to highway investments in Chapter 3, then considering critical economic and social infrastructure features. In Chapter 4, we extend our economic and social return rates by discounting the harmful environmental impacts of roads and computing a novel and regionalized sustainable and equitable return rate for highway investments. From these three empirical exercises, we apply economic, social, and environmental criteria and clustering methods to identify potential prioritization areas for road interventions in Brazil.

First, we describe the economic, social (equality-related), and environmental components of the return rates to highway investments calculated in the past Chapters of this Thesis. By doing so, we clarify the different factors that increase (decrease) the return rates for each Brazilian Immediate Geographical Region (RGI). Second, we propose eligibility and prioritization standards for road policies at the regional scale. The eligibility criteria guarantee economically attractive returns while respecting social and environmental issues. The prioritization criteria classify RGIs by considering their economic profitability and the propensity to reduce regional inequalities and avoid environmental damage from roads. Those standards constitute original inputs to develop focalized road policies. Third, we extend our analysis by considering more aggregated spatial levels. We apply clustering methods to create highway policy zones, minimizing the dissimilarities between regions in economic, social, and environmental road-related issues. Thus, we provide novel evidence for planning and evaluating national road policies, especially those crossing extensive geographical areas and different regions. Finally, we carry out an *ex-post* evaluation of the "old" PAC (PAC 1 and 2, 2007-2018), raising some ways to use our findings and prioritization criteria in different spatial scales. To the best of our knowledge, this is the first *ex-post* impact evaluation of the PAC discussing its priorities in economic, social, and environmental terms. In this way, we contribute to the specialized literature on transportation infrastructure and regional development (Banerjee *et al.*, 2020; Baum-Snow *et al.*, 2020; Duranton *et al.*, 2014; Faber, 2014; Herzog, 2021; Jaworski and Kitchens, 2019; Lu *et al.*, 2022; Zhang and Ji, 2019; Zhang *et al.*, 2020) by providing new evidence to develop more cost-effective, inclusive, and sustainable road policies at diverse spatial scales, which we expect to be applicable to interventions of different extensions as well as reaching diverse regions.

The remainder of this paper is organized as follows. Section 2 describes the economic, social, and environmental components of the return rate to highway investments. Section 3 details the priority classification for highway investments. Section 4 outlines highway investment zones for the national policy analysis. Section 5 presents an *ex*-post evaluation of the PAC, looking at economic, social, and environmental issues in highway investments. Section 6 concludes.

5.2. The economic, social, and environmental components of the Sustainable and Equitable Return Rate to Highway Investments (SERR)

To develop the eligibility and prioritization criteria for road policies, we start from our preferred and broader return rate to highway investments calculated in Chapter 4. Our Sustainable and Equitable Return Rate to highway investments (SERR) is calculated considering economic issues – through the road impacts on productivity –, social features – at reducing infrastructure-related inequalities across municipalities –, and environmental aspects – at increasing GHG emissions. Then, we decompose the SERR into three components as follows:

$$SERR_r = EC_r + SC_r - GEC_r \quad (1)$$

Where EC is the economic component, SC is the social (equality-related) component, GEC is the environmental component linked to GHG emissions, and *r* represents the 510 Brazilian RGIs.

To obtain the SC, we subtract the RR Average (RR) from the RR Efficient & Road Specialized & Redistributive & Equative (RREE), both computed in Chapter 3, as follows:

$$SC_r = RREE_r - RR_r \quad (2)$$

The RR only considers the economic returns of road investments, ignoring heterogenous impacts in social terms and disregarding environmental effects from roads. The RREE captures heterogenous road impacts on productivity depending on infrastructure features such as efficiency, road specialization, redistribution, and equity. Redistribution is the policy purpose that uses road interventions to foster regionally balanced economic growth by targeting poorer localities. Similarly, equity means investing in places characterized by low infrastructure endowment, equalizing the territory. In Chapter 3, we found a larger road impact on GDP *per capita* in less developed municipalities characterized by poor infrastructure and low productivity, suggesting a "social bonus" at investing in roads in those localities. In this sense, the RREE includes economic and social (equality-related) issues, and the SC captures the increased (decreased) road impact (the "social bonus") for places with lower GDP *per capita* and road endowment levels.

Consequently, the economic component is represented as follows:

$$EC_r = RREE_r - SC_r \quad (3)$$

The EC is the RREE discounted from its social component and obviously equalizes the RR from Chapter 3. In other words, the EC represents the isolated economic returns of road investments, the road impact on productivity overlooking heterogeneous effects from redistribution and equity, or even the road impact on productivity supposing the same average road impact for all regions.

Finally, the GEC is equal to the CO2 emissions Return Rate (ERR) in Chapter 4, which considers the road impacts on GHG emissions related to the energy sector. As highway investments increase GHG emissions, the GEC enters with a negative sign in Equation 1. We avoid the CO2 emissions from the land use change and agriculture sectors to direct our public policy recommendations to the transportation authorities. In addition, we impose an additional environmental condition to highlight the potential expanding effects of roads on land use change GHG emissions as follows:

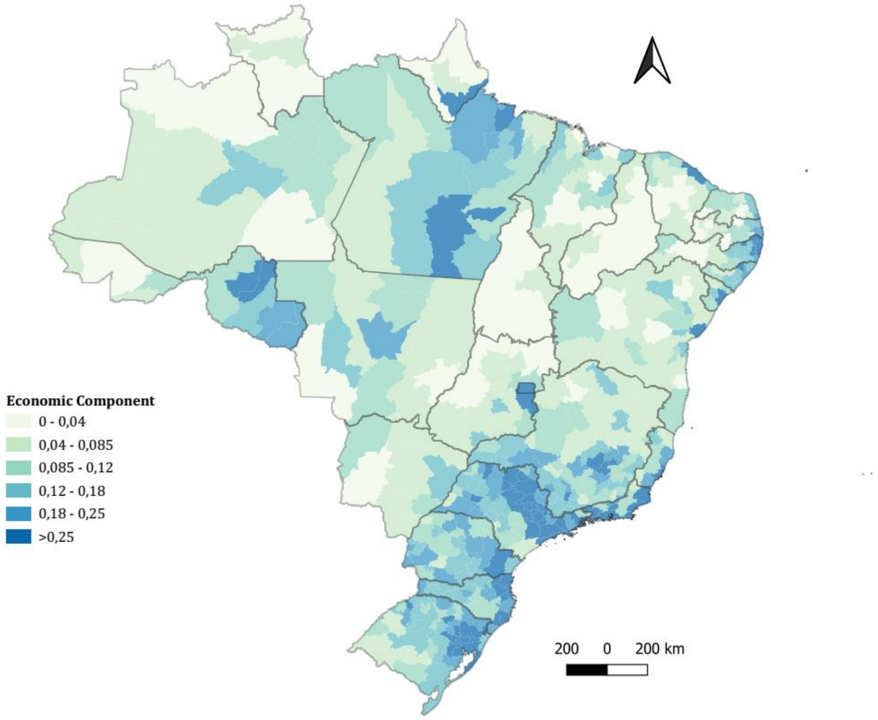
$$Environmental\ focal\ points \begin{Bmatrix} 1\ if\ SERR_r^{energy} \geq 0.12\ and\ SERR_r^{all} < 0.12\ and\ GEC_r^{Decile} = 10 \\ 0\ otherwise \end{Bmatrix} \tag{4}$$

Where $SERR_r^{energy}$ represents the return rate discounting the road impacts on the energy sector CO2 emissions, $SERR_r^{all}$ is the return rate deducting the highway impacts on the CO2 emissions of all sectors including land use change, and GEC_r^{Decile} is the GEC decile. Equation 4 indicates that we have an environmental point of attention when the RGI attends three requirements. First, the return rate to highway investment is higher than 12% considering energy-related CO2 emissions. Second, the SERR drops to levels smaller than 12% at taking all GHG emission sectors, including the land use change and agriculture. Third, the RGIs are in the top 10% of the GEC variable and assume large values (above 7.24%) for the environmental component. Therefore, the environmental focal points are regions wherein the environmental component is critical in defining road profitability, and broad government participation is expected in both transportation and environmental sides to ensure that the environmental costs do not offset economic and social benefits from roads.

Figure 5.1 shows the SERR components as well as the SERR itself. The RGIs with higher economic returns are concentrated in the more developed regions of the country – the South and Southeast –and in some inland areas, wherein the agriculture sector has boosted the economy in the last decades. Those more profitable places coincide largely with regions crossed by privately managed highways, suggesting the higher attractiveness of those roads for private partners mainly guided by economic profit. Figure A1 in Appendix A shows the privately managed highways and the economic component, supporting our arguments. In Table A1 in Appendix A, we summarize the SERR and its components by public or private roads, suggesting that private partners are likely looking for road segments with higher expected economic returns.

Figure 5.1. The Sustainable and Equitable Return Rate to highway investments (SERR): economic component (a), social component (b), environmental component (c), and SERR (d)

(a)



(b)

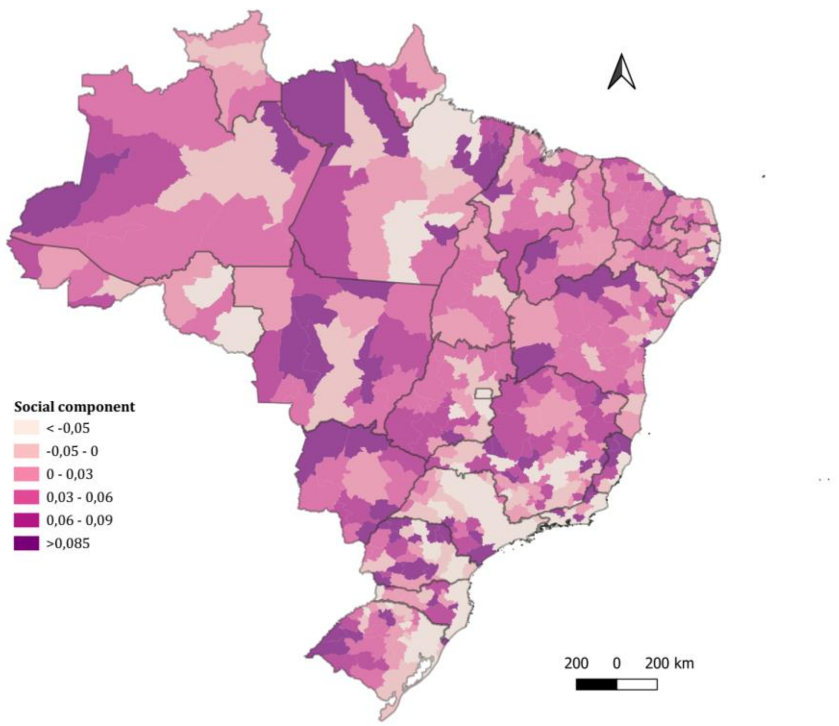
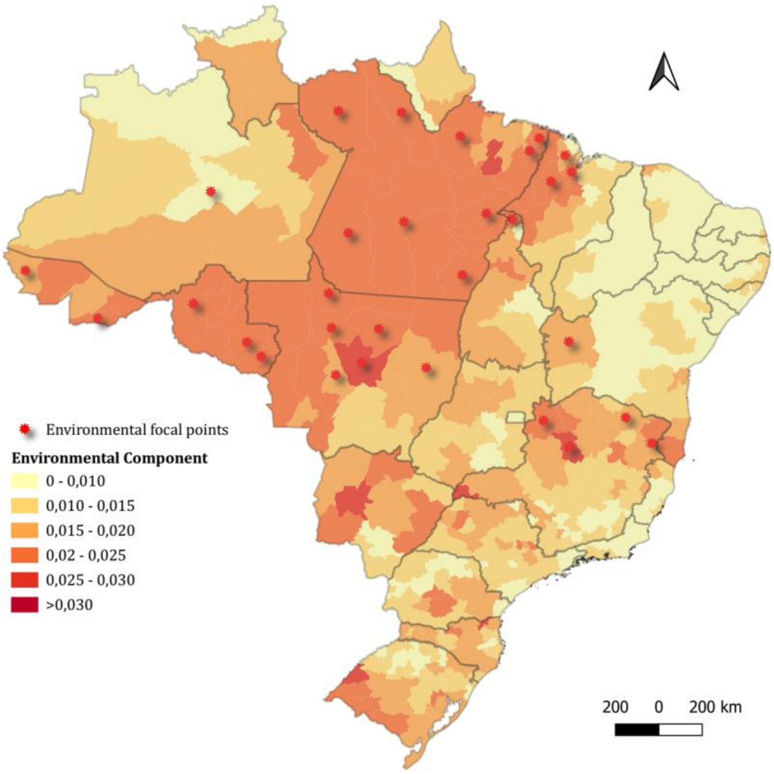
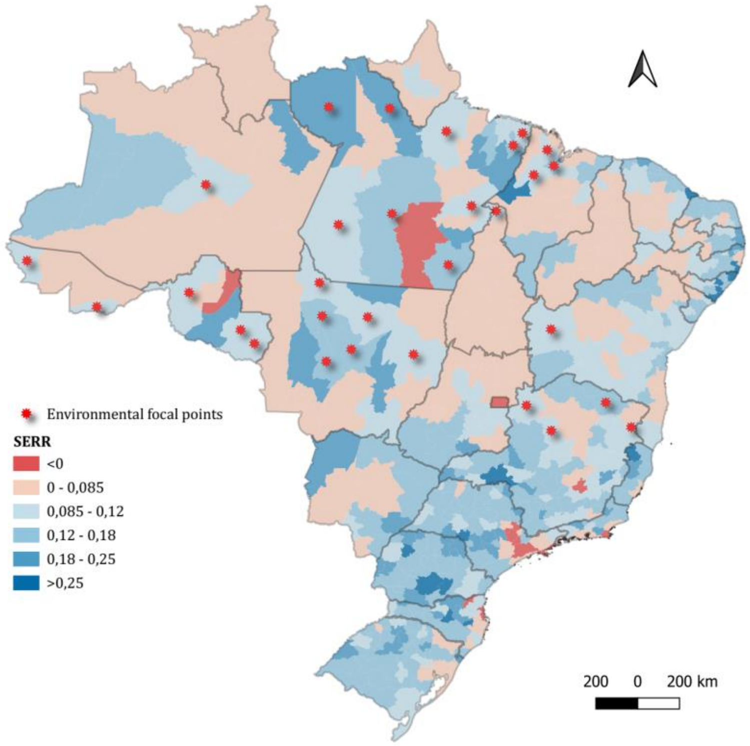


Figure 5.1. The Sustainable and Equitable Return Rate to highway investments (SERR): economic component (a), social component (b), environmental component (c), and SERR (d)

(c)



(d)



Source: authors' elaboration.

The social component tends to be larger for the poorer and more remote RGIs and smaller for the wealthiest regions. The results in Chapter 3 pointed out that road impacts on the economy are higher for less developed and poorer infrastructure-endowed places. Thus, investing in highway development might be an important inclusive policy tool. This can be seen in the strong social component in the Northeast and North regions, which are largely characterized by low productivity and transportation infrastructure levels.

The environmental component is more significant in the North and part of the Mid-West regions. As we argued in Chapter 4, this result is closely related to the recent deforestation in the Brazilian Amazon, places that still have huge forest areas and are suffering the most from the environmentally damaging effects of roads. The environmental focal points presented in the map corroborate the findings. In these cases, the interaction between transportation and environmental institutions is critical to guarantee minimum negative road impacts on the environment.

Regarding the SERR in Figure 5.1 (d), we can observe several profitable places across the country. This result indicates the well-known need to develop the Brazilian transportation infrastructure sector. There are a relevant number of profitable RGIs in the poorer Northeast and North regions as well as in the richer South, Southeast, and Mid-West, suggesting that road policies might be used for national and regional development purposes.

Finally, we analyze the correlation between the SERR and its components with road features such as efficiency, redistribution, equity, and road specialization. Figure A2 in Appendix A shows the correlation matrix. The economic component positively correlates with all road features, suggesting that regions with higher efficiency, GDP *per capita,* road endowment and transportation reliance are more profitable for highway investments. On the other hand, the social component is negatively correlated with all those variables, indicating less developed regions as those benefiting more from the "*social bonuses*" of road interventions. The environmental component presents a strong and negative correlation with equity while a positive correlation with road specialization. These results suggest that regions highly specialized in roads present higher environmental costs from roads. However, as these highways densify in space – as the equity measure indicates – the environmental cost becomes smaller. Next, we set the eligibility and prioritization criteria.

5.3. The priority classification for highway investments

In this section, we provide eligibility and prioritization criteria for highway policies in Brazil. The eligibility requirement follows a minimum profitability cut-off that considers social and environmental issues. The prioritization criteria classify the RGIs by their profitability, amplifying redistribution and equity, and minimizing environmental damage.

5.3.1. Eligibility criteria

The eligibility standard seeks to guarantee a minimum level of profitability to road investments considering social and environmental conditions. To do this, we first take the SERR calculated in Chapter 4, as this return rate includes all highway investment components. Then, we compare the SERR with two benchmark rates. The first is the Social Discount Rate (TSD)(Ministry of Economy, 2021) of 8.5% per year. The TSD is broadly used to evaluate infrastructure project returns in Brazil. The second one is the basic interest rate (Selic rate), which guides most private investment returns in the country. The Selic rate was around 12% per year in the past twelve months, the value we take as the second cut-off. While the TSD is more suitable for evaluating public infrastructure investments, the Selic rate is expected to capture the minimum profitability level expected by the private sector more realistically.

As the Brazilian economy has presented several financing and funding issues over its history (Armijo and Rhodes, 2017; Burrier, 2019; Carranza *et al.*, 2014), it would be unreasonable to establish eligible RGIs with return rates below the cut-off rates. In addition, we want to provide policy recommendations in the light of an already started road program (the "New" PAC), which drives us to suggest interventions that might be effective even in the short and medium term. We are not arguing that the non-eligible regions must not receive highway investments, but that infrastructure projects in those regions may lack profitability or present issues related to increasing inequality or environmental damage. Then, they might demand more time and effort in the planning and project design phases. On the other hand, the eligible RGIs are *win-win* localities wherein profitability is achieved even considering potential regional inequalities and environmental harm from roads.

It is important to note that the SERR includes social and environmental features. In addition, the SERR assumes value zero for some highly road-efficient places wherein inefficiencies are expected due to huge geographic, environmental, expropriation, and interferences infrastructure project costs. Those places will also need highway investments, but they might need more careful planning, design, and execution from the government. Due to its large expected economic return, private investment might emerge as a potential solution, as it works for the state roads in the State of São Paulo, and road investments in maintenance might be critical for areas with extensive infrastructure stock.

In this context, the eligibility criteria can be described as follows:

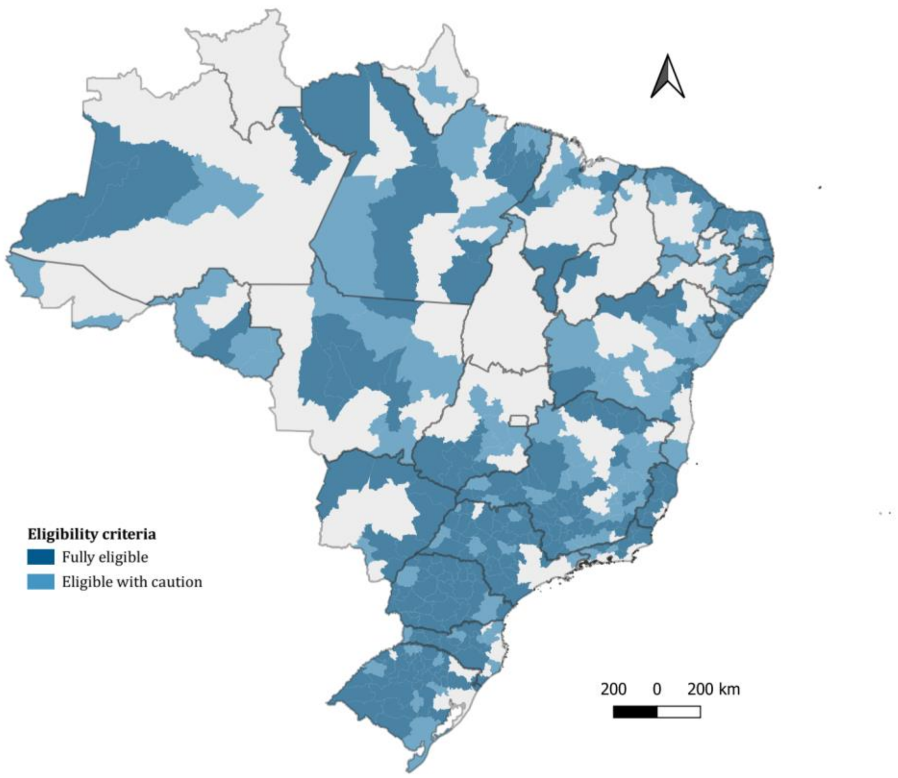
$$Eligibility\ criteria \left\{ \begin{array}{c} Not\ eligible\ if\ SERR_r < 8.5\% \\ Eligible\ with\ caution\ if\ SERR_r \geq 8.5\%\ and\ SERR_r < 12.0\% \\ Fully\ eligible\ if\ SERR_r \geq 12.0\% \end{array} \right\} \quad (5)$$

The eligibility criteria ensure the return rate to highway investments to be larger than a minimum profitability level (8.5%). For those eligible places assuming

SERR values between 8.5% and 12%, we suggest caution as profitability might be insufficient, especially for policies aimed at attracting private resources. RGIs presenting SERR greater than 12% are deemed fully eligible for road investments. When the SERR is smaller than the minimum 8.5% cut-off rate, we consider the RGI not eligible.

Figure 5.2 shows the eligible RGIs. There are 233 (45.69%) fully eligible RGIS, 118 (23.14%) eligible with caution RGIs, and 159 (31.18%) not eligible RGIs. In other words, more than 68% of the RGIs seem to demand highway investments, even considering social and environmental issues. In the next step, we classify those eligible RGIs to provide the best options for fostering economic activity, reducing inequalities, and restricting environmental damages.

Figure 5.2. Eligible RGIs for highway investments



Source: authors' elaboration.

5.3.2. The prioritization criteria

The prioritization requirement aims to classify the RGIs according to the three SERR components. The main goal of this proposal is to allow policymakers to target places where highway investments might promote more sustainable and regionally balanced economic growth. We start by taking the eligible RGIs.

The priority criteria are defined in Table 5.1. Priority 1 is the best-case scenario, composed of profitable RGIs with EC and SC values above the median and

GEC values below the median. In other words, Priority 1 regions are those wherein *win-win* combinations are fully achieved, configuring highway investments potentially increasing economic growth, reducing inequalities, and minimizing environmental costs. There are 28 RGIs classified as Priority 1, showing a high return rate to highway investments (Columns Average SERR) of 21.5%, three to four times larger than the average for the group of non-eligible RGIs.

Table 5.1. The prioritization criteria: RGI level

Prioritization level	EC	SC	GEC	RGIs (≥12.0%)	RGIs (≥8.5%)	Average SERR (≥12.0%)	Average SERR (≥8.5%)
Priority 1	Above median	Above median	Below median	28	28	0.215	0.215
Priority 2	Below median	Above median	Below median	38	71	0.160	0.133
Priority 3	Above median	Above median	Above median	34	34	0.210	0.210
Priority 4	Above median	Below median	Below median	40	62	0.144	0.129
Priority 5	Below median	Above median	Above median	37	58	0.155	0.135
Priority 6	Below median	Below median	Below median	2	14	0.124	0.103
Priority 7	Above median	Below median	Above median	54	78	0.154	0.139
Priority 8	Below median	Below median	Above median	-	6	-	0.098
Not eligible	-	-	-	277	159	0.074	0.053

Source: authors' elaboration.

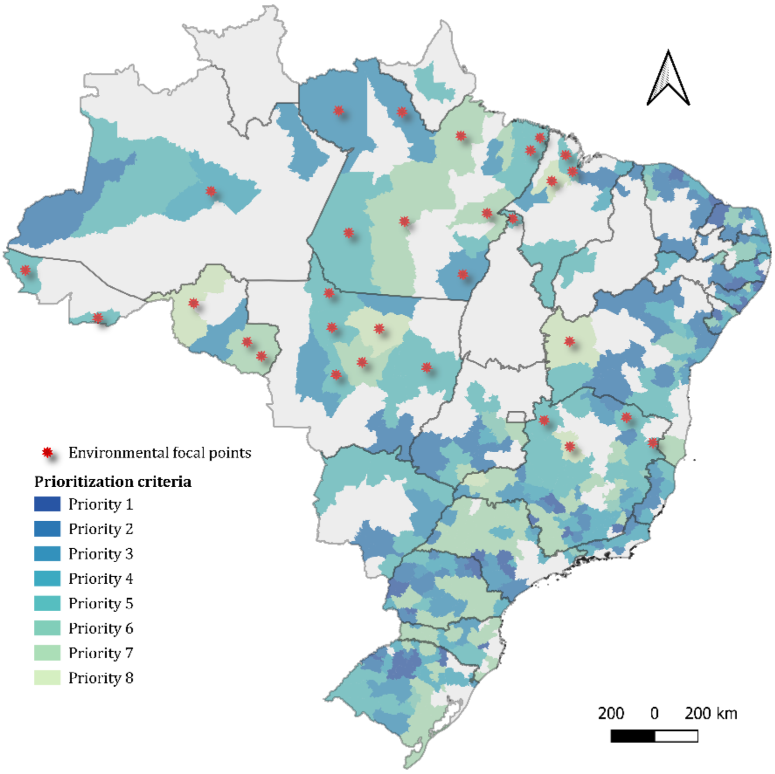
Priority 2 covers those RGIs with suitable social and environmental conditions but presenting EC values below the median. In this case, we can observe a smaller SERR average compared to Priority 1. Nonetheless, the average SERR of 16% and 13.3%, considering the cut-off rates of 12% and 8.5%, respectively, suggest a high profitability for those localities.

Equivalent interpretations can be made to the following priority criteria. Priority 3 meets the EC and SC requirements but presents large environmental costs. Priority 4 attends economic and environmental criteria, while the SC is smaller than the median. Priority 5 does not meet EC and GEC requirements but does so for the SC. Priority 6 satisfies environmental conditions but presents values below the median of EC and SC. Priority 7 respects the EC requirement, while SC and GEC assume inappropriate values. Finally, Priority 8 is those RGIs with SERR above the cut-off of 8.5%, but that do not meet any other prioritization criteria.

Figure 5.3 shows the results. Regarding Priority 1, we can observe some best-case scenarios for road investments in the coastal area of the Northeast region and inland RGIs in the Southeast and South regions. When we include Priorities 2 to 4 into the analysis, which is characterized by not attending one of the prioritization requirements, we can identify some potential groups for road investments. The

Priority RGIs in the Northeast coastal zones are enlarged, indicating that the region is a critical target for fostering sustainable and inclusive development through road improvement. Some Priorities also emerge in inland areas of the Northeast region and the north of the State of Minas Gerais, places characterized by low levels of GDP *per capita* and infrastructure endowment. By extending the Priority levels, considerable Priority RGIs appear in the coastal and inland areas of the Southeast region, localities marked by higher productivity and population density levels and high demand for infrastructure interventions. The same pattern holds for the new Priority RGIs in the inland areas of the South and Mid-West regions, distinguished by agribusiness as an economic driver.

Figure 5.3. Priority RGIs for sustainable and equative highway investments



Source: authors' elaboration.

Finally, we extend our examination by looking at Priorities 5 to 8. Those RGIs do not meet two or more prioritization criteria but are profitable. Several Priority RGIs arise in some of the wealthiest areas of the country. This result is expected as we do not meet social and environmental requirements in many regions. For instance, the novel Priority RGIs in the more developed Southeast and South regions present a low social component. In addition, several environmental focal points emerge as Priority RGIs in the Mid-West and North regions. In these cases, the economic component mainly guides the priority classification, likely due to the expanding agricultural activity and the consequent need for infrastructure development. However, this economic benefit is being constrained by huge environmental costs from transportation activity, deforestation, and small social bonuses from reducing inequalities. That is why those RGIs do not appear to be the first priority in our rank.

The prioritization criteria allow identifying RGIs wherein highway investments are expected to foster economic growth, alleviate inequalities, and prevent environmental costs. As eligible RGIs fall in the priority classification, the lower the predicted road investment returns become and the greater the public policy efforts to prevent social and environmental harm. In the next section, we complement this more disaggregated analysis at the RGI level by evaluating the return rate to road investments at larger regional scales. We propose some highway investment zones for road policies by aggregating RGIs with similar SERR components.

5.4. Creating highway investment zones for national policy analysis

We have classified the RGIs in priority terms for road policies. This input might guide policymakers to invest in places to foster economic growth, reduce inequalities, and avoid environmental damage from road development. However, some practical issues may emerge. Some priority RGIs might be geographically isolated, and a place-based investment would be unrealistic. For instance, we do not expect the federal government to allocate small road segments (less than 100 km) to concession programs even if the economic, social, and environmental returns are large enough in that RGI and its surroundings. For instance, the average highway length of federal concessions is 522 km. In addition, national road interventions are likely to target specific road segments linking important economic centers to other economic zones, highly populated regions, ports, and so forth. In several cases, those road investments will cover extensive highway segments, and a more aggregated view might provide some valuable additional insights into road policies. For instance, those points are not problems when evaluating a bridge building or a road segment upgrade in an urban area. Then, the more regionally aggregated proposal complements the RGI scale as we have local, regional, and national road interventions.

In this section, we complement Sections 2 and 3 by aggregating our findings using two different approaches. First, we evaluate our SERR and its components at the state level. It is important to clarify to state governments and society how national highway investments might be a critical policy tool to stimulate sustainable and balanced economic development in each state. In addition, this aggregation might provide insights for state level road policies as well. Second, we use spatial clustering methods to generate highway investment areas based on the three SERR components. In this exercise, we seek to overcome administrative borders by providing road policy zones based on the similarity between the RGIs in terms of economic, social, and environmental road issues constrained by spatial proximity.

From this, we provide some contributions to think about inclusive and sustainable public road policies at the national level from a regionalized perspective.

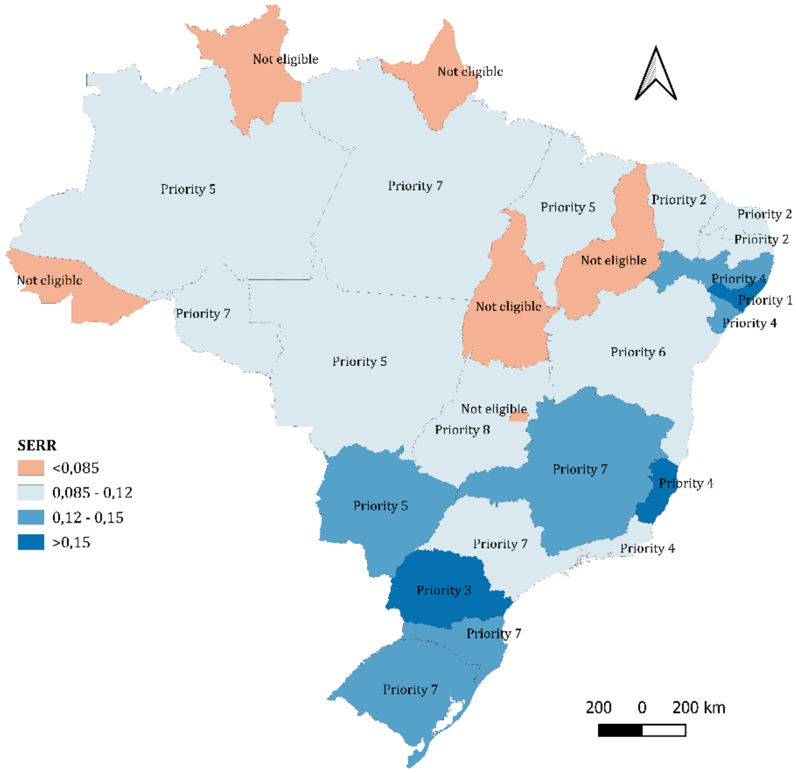
5.4.1. State-level SERR and components

In this exercise, we aggregate the RGI data used so far at the state level. We do this by taking simple averages of the EC, SC, GEC, and SERR variables for each of the 26 Brazilian states plus the Federal District. Next, we apply the same prioritization criteria following Table 5.1. Figure 5.4 shows the state SERRs. The economic, social, and environmental components can be seen in Figures B1, B2, and B3 in Appendix B. Most (77.8%) of the states meet the eligibility criteria, being profitable for highway investments.

It is interesting to note that the states with Priority 1 and 2 are localized in the Northeast region. Even taking a more aggregated SERR measure, our findings continue to suggest that investing in those places might promote a more sustainable and inclusive economic growth. On the other hand, while the most profitable states are concentrated in the South and Southeast regions, they lack social or environmental requirements and are mostly classified with lower priority levels (mainly 7, but also 3 and 4). Finally, most Mid-West and North states present SERR between 8.5% and 12%, suggesting some caution in proposing road policies in those places. This is especially due to the higher environmental costs in those states.

Our findings provide some contributions to public policies. First, we can project the expected economic benefit from road investments, respecting social and environmental issues. This is useful for federal and state governments in designing national, regional, and local road interventions. Second, the prioritization criteria provide a practical framework for structuring public road policies according to society's concerns regarding economic growth, reducing inequalities, and dodging environmental damage. At the state level aggregation, we observe only one state (Alagoas) satisfying the three component requirements, i.e., presenting profitability higher than 8.5%, high social and economic components, and low environmental costs. Except for this case, the Brazilian federal government must decide to deal with issues in at least one of the three SERR components. In this context, the prioritization criteria provide a tool for weighing the component's importance when developing road policies across the country. Moreover, Figure 5.4 gives us some understanding of the inevitability of articulating a broad transportation policy, including both federal and state-level roads. If states with high social and environmental priority lack resources to invest in infrastructure, we might expect increasing inequality and environmental degradation from roads. The same holds whether more developed states are implementing strong road programs, as it occurs in São Paulo and other high-income states in the South, Southeast, and Mid-West regions.

Figure 5.4. SERR and prioritization criteria: State-level aggregation



Source: authors' elaboration.

In this application, we focus on administrative borders. Nonetheless, places in different states might present similar road features, and geographically bounding our analysis and policy implications into state boundaries can hinder the development of an effective national road program. In the next section, we seek to overcome this issue by establishing road policy zones based on spatial clustering methods.

5.4.2. The spatial clustering framework

In this section, we create highway investment zones for policy analysis. To this end, we apply the *Spatial 'K'luster Analysis* (SKATER)(Assunção *et al.*, 2006), an efficient data-management technique for finding homogeneous groups of elements in a heterogeneous dataset (da Silva *et al.*, 2014). This method allows us to consider groups with a spatial location, in our case, the RGIs.

Infrastructure development is spatial by nature, as highways will serve specific geographic areas connecting one place to another (Straub, 2011). In other words, road policies are designed spatially, and they will cover road segments aimed at impacting specific localities, and the effectiveness of this instrument is deeply correlated with the infrastructure characteristics of regions. For instance, we expect more developed regions to be more attractive for private investments due to their higher projected profitability. Regions presenting vital environmental issues will likely require additional administrative and economic efforts as environmental licenses and financial resources, and places characterized by lacking infrastructure endowment and lower income levels may not be financially attractive, which will probably raise the importance of public investment in developing those regions. Therefore, clustering regions in terms of economic, social, and environmental road-related issues might be interesting for re-thinking national road policies, which is our goal in this exercise.

5.4.3. Method

The SKATER builds off a connectivity graph to represent spatial relationships between neighboring regions, where each region is represented by a node and edges represent connections between them. Edge costs are calculated by evaluating the dissimilarity between neighboring areas. The connectivity graph is reduced by pruning edges with higher dissimilarity until we are left with *n* nodes and *n-1* edges. At this point, any further pruning would create subgraphs, and these subgraphs become cluster candidates. The SKATER algorithm works by iteratively partitioning the graph by identifying which edge to remove to minimize the within-cluster sum of squares (a measure of how tight each cluster is) and maximizing the between-cluster sum of squares (a measure of how separated each cluster is from the others). (Assunção *et al.*, 2006).

We use the economic, social, and environmental SERR components as our interest variables in the clustering process. We construct the spatial constraints using a first-order queen matrix of spatial weights. In addition, we try geographic area constraints to guarantee a minimum cluster size. This constraint prevents highway policy zones (clusters) from being too small, which might be an issue in designing public road interventions and concession programs in several circumstances.

We apply the SKATER method, trying several combinations of the number of clusters and geographic area constraints. The results in terms of the total within-cluster sum of squares, the between-cluster sum of squares, and the ratio of between to total sum of squares, as well as a detailed description of the SKATER application, can be seen in Appendix C. Based on those statistics and the practical applicability of the clusters, our preferred SKATER result consists of 10 highway policy zones constrained by having at least 3% of the national geographic area. As a robustness check, we also use the combination of 27 clusters restricted by having at least 1% of the national geographic area. Those clusters complement one another due to the hierarchical nature of the SKATER method, in which the new and smaller clusters (in the 27-cluster scenario) are partitions of the previous and larger clusters (in the 10-cluster scenario).

5.4.4. Results

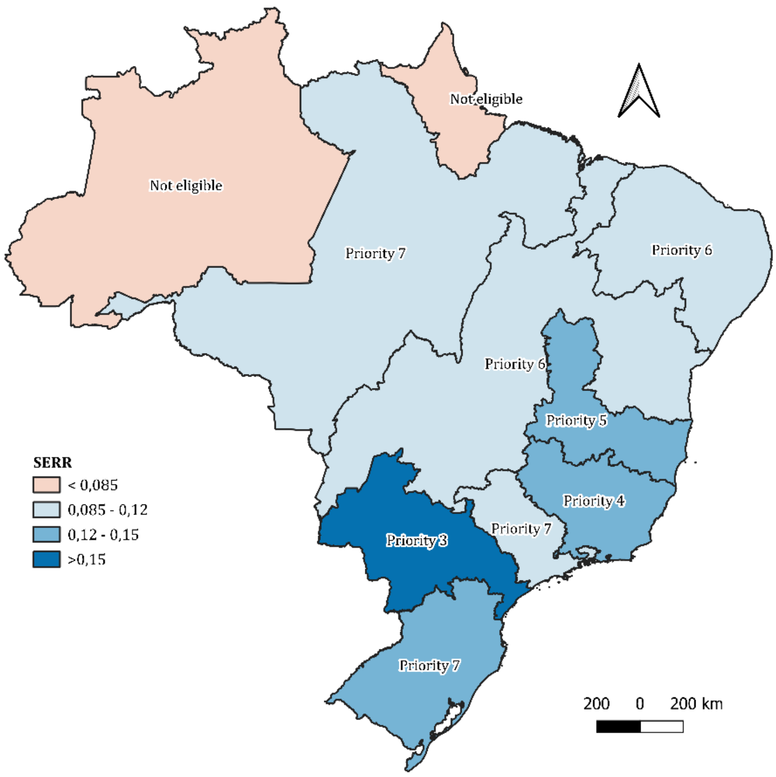
Figure 5.5 shows the results of our preferred highway policy zones. The map includes the SERR values as polygon colors and indicates the priority level of each zone. Eight out of ten zones present SERR values higher than the 8.5% cut-off, suggesting the high road profitability in the country in this new aggregation level as well. As we can see by taking the zone aggregation, there is no region in the best-case scenario (Priority 1) or even in Priority 2. This result suggests that spatially broader road policies will likely deal with one or more economic, social, or environmental issues in road planning, designing, and execution. In addition, by aggregating RGIs in larger zones, our findings provide some elucidation on the profitability of road programs covering extensive geographical areas and road segments. When evaluating the more disaggregated RGI level, we could find several regions in Priority 1. However, that is not the case for the SKATER zone level, suggesting that the zones mix profitable and non-profitable RGIs. The SERR values are more concentrated in the zone scale, presenting a minimum value of 0.058 and a maximum value of 0.154. This result raises confidence in our cluster results in representing the Brazilian road infrastructure sector, especially in planning and designing national public policies seeking to connect distant locations or improve the connections among them. In the Brazilian context, it seems more reasonable and representative to expect a more moderate return on highway investments when they cross extensive regions with different levels of income per capita, infrastructure endowment, and geographical, environmental, and human-physical costs. This greater complexity of extensive buildings is often followed by inefficiencies due to institutional weaknesses in the country (Amann *et al.*, 2016; Armijo and Rhodes, 2017; Burrier, 2019; Raiser *et al.*, 2017).

The average highway policy zone SERR is 0.110, which is in line with the internal return rates (IRR) between 8% and 10% established by the National Land Transport Agency (ANTT) to guide several concession programs across the country, including road segments over 1,000 km in length. Those findings are particularly interesting in the infrastructure sector due to its spatial nature, wherein extensive road buildings and concession programs should cover places in different income *per capita* and infrastructure endowment levels to assure poorer and non-profitable regions receiving road investments coming from the profits in more lucrative areas. This implicit cross-subsidization between regions might act as a reducing inequality policy tool through road investments (Sousa and Da Silva Filho, 2022). By mixing profitable and non-profitable regions but respecting some degree of similarity between them in economic, social, and environmental terms, our spatial clustering application seems to consider this infrastructure sector particularity indirectly.

Next, we evaluate the SERR components and the efficiency, redistribution, equity, and road specialization features at the highway policy zone level. Table 5.2 exhibits the results. The economic component is larger in the South and Southeast regions, represented by zones 3, 4, 5, and especially 7, the latter assuming an EC of 30%. In general, those regions are marked by higher GDP *per capita*, infrastructure endowment, road specialization, and efficiency, which explains, to a large extent, the higher expected economic return from road investments there.

Figure 5.5. Highway policy zones description (a), and SERR and prioritization criteria (b)

(a)



(b)

Source: authors' elaboration.

Table 5.2. EC, SC, GEC, SEER, and road features: Highway Policy Zones

Zone	EC	SC	GEC	SERR	Priority	Efficiency (GDP/Road Stock)	Redistribution (GDP *per capita)*	Equity (Road Stock/Area)	Road specialization (φ)
1	0.098	0.023	**0.007**	0.114	6	9.644	**14.690**	8.122	0.038
2	0.077	0.028	0.012	0.093	6	9.427	26.770	5.804	0.040
3	0.190	-0.035	0.014	0.141	7	20.192	43.735	7.904	0.042
4	0.149	-0.017	0.010	0.122	4	21.251	30.115	9.124	0.042
5	0.126	0.041	0.013	**0.154**	3	11.392	39.532	7.987	0.042
6	0.122	0.013	**0.022**	0.113	7	30.534	27.980	2.658	0.043
7	**0.300**	**-0.176**	0.014	0.110	7	**66.469**	**49.596**	**11.407**	**0.044**
8	**0.056**	0.036	0.016	0.076	Not eligible	7.214	15.817	**0.685**	**0.032**
9	0.085	**0.056**	0.018	0.122	5	**6.541**	17.006	3.201	**0.044**
10	0.152	-0.079	0.015	**0.058**	Not eligible	15.883	19.497	1.351	0.042

Source: authors' elaboration.

Regarding the social component, zones 9 and 5 present the higher values. Region 9 seems to be a redistributive priority due to its low GDP per capita and road endowment levels. Zone 5 is distinguished by a high GDP per capita level but an average value for the equity variable, suggesting that roads might be an essential policy tool to expand the region's transportation connectivity. On the other hand, zone 7 presents a huge harmful social component due to its large values for all the infrastructure characteristics.

Zone 6 has a higher environmental component. Most of this zone is in the Brazilian Amazon, which presents extensive forest areas. Constructing roads in those places has directly promoted the spread of transportation activity but also increased deforestation. The result is a higher environmental cost from direct and indirect GHG emissions, which seems to be well captured by our clusters.

In short, the zones constructed by applying the SKATER method seem to suitably represent the Brazilian infrastructure road sector. These areas are particularly appropriate for evaluating national infrastructure policies wherein highway segments cross large geographic areas, as the clusters aggregate similar regions in economic, social, and environmental terms, respecting spatial proximity. In addition, the highway policy zones might be combined with the more disaggregated state and RGI levels, spatial scales that can complement one another in planning and designing national, regional, and local road policies.

5.5. An *ex-post* evaluation of the "old" PAC (2007-2018): looking at economic, social, and environmental issues in highway investments

In this section, we provide some paths to evaluate road programs in the light of our calculated economic, social, and environmental components. We intend to provide practical tools to allow stakeholders to apply the methodology built up throughout this thesis when evaluating infrastructure policies in Brazil.

To that end, we start by running an *ex-post* impact evaluation of the PAC by comparing its results with what would have happened if the interventions had been allocated randomly across regions. We divide this naive exercise into some steps. First, we reassess the PAC highway investment data constructed in Chapter 2 and aggregate it at the RGI level. To identify the regions treated by the PAC, we use a dummy variable assuming value one if the RGI received some road investment and zero otherwise, which we call treatment in this application. Using the treatment variable instead of the PAC highway investments reduces the problem of randomly placing high investment values to locations where building a highway may not be reasonable. The PAC targeted 215 out of 510 RGIs. Second, we randomly assign our treatment variable to regions, repeating the process 10,000 times to ensure a suitable number of observations. From this, we provide a benchmark for what would have occurred if road investments had been allocated accidentally, regardless of economic, social, environmental, and political issues. Third, we compare the PAC's EC, SC, GEC, and SERR values with those provided by the simulations.

Figure 5.6 shows the results. Findings suggest the program being economically biased, allocating investments to RGIs with a higher economic component, corroborating previous results in Chapters 2 and 3. The average simulation results return an EC of 0.136, while the average EC for the regions treated by the PAC is 0.145. On the other hand, the social component of -0.015 is smaller than the average simulation value of -0.0058. This finding supports the argument that the PAC assigned a considerable amount of financial resources to more developed regions wherein road improvements did not play a predominant role in raising productivity during the evaluated period, lacking some part of the “social bonuses” expected at investing in poorer places. In addition, wealthier regions are characterized by higher geographic, environmental, expropriation, and interference infrastructure project costs, which appears to have narrowed the highway investments' impact on the economy due to inefficiencies in the planning, design, and execution phases. The GEC is slightly higher for the PAC regions (0.014) in comparison with the simulation average (0.013). As a robustness check, we also run simulations considering the road impacts on GHG emissions of all sectors. Results are exhibited in Figure D1 in Appendix D. Based on this exercise, the environmental PAC costs rise to 0.033, while the average simulation GEC considering all sectors is 0.030. Whereas the difference between PAC and simulation values is still small, those values more than double in relation to the GEC considering the energy sector only, reinforcing the importance of complementary policies to secure environmental preservation.

By evaluating all SERR components together, our results indicate that the PAC achieved a similar outcome even compared with average values of RGIs wherein economic, social, environmental, and political issues are expected to be random. This is represented by a PAC SERR of 0.116, while the simulation average is 0.117. The PAC SERR falls to 0.096 when considering the road impacts on GHG emissions in all sectors, including land use change, whereas the simulation average is 0.100. While the PAC yielded meaningful economic effects on the economy, the program targeting criteria did not seem effective regarding social and environmental components, assimilating itself to a policy implemented at random.

Next, we run the same randomization process but allocating roads to the eligible RGIs following our framework in Section 5.3. The results are described in Figure 5.7. The average simulation values for the EC, SC, GEC, and SERR are then 0.140, 0.025, 0.012, and 0.152, respectively. Even considering eligible RGIs, which present highway returns above 8.5%, the average EC for PAC RGIs is higher than the simulation average. Nonetheless, the PAC social and environmental results are far behind from which could be achieved by applying the resources to prioritized regions. This result suggests that the above-average economic benefits from the PAC were somewhat offset by social and environmental losses.

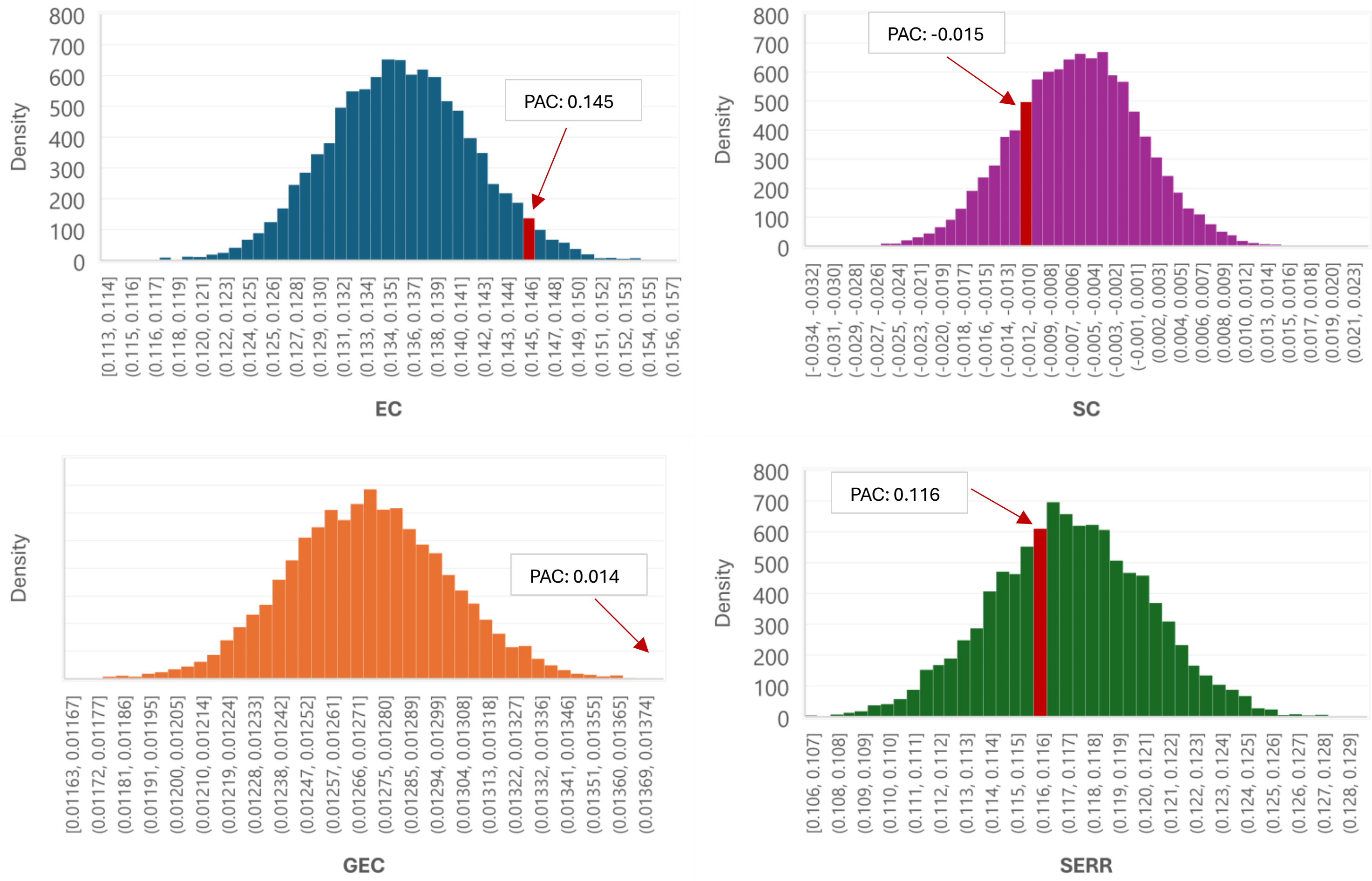
In Figure 5.7, we cannot visualize the PAC values for the SC, GEC, and SERR, indicating that the social and environmental PAC outcomes would barely be achieved even under 10,000 random combinations of targeted priority RGIs. It is important to note that the PAC values will be even more distant from the simulation values when selecting only higher priority regions (especially Priority 1, but also Priorities 2 to 4).

As complementary tests, we replicate the same simulations at the state and highway policy zone levels but using the highway investment values as all states and zones were treated by the program. Results can be seen in Table D1 in Appendix D, corroborating previous findings using the RGI spatial scale. In short, the PAC road allocation provides poorer social and environmental results when compared with the randomization averages, while it provides superior economic returns. At the state level, we calculate the EC, SC, GEC, and SERR for the PAC being equal to 0.150, -0.017, 0.014, and 0.120, respectively. Findings are corroborated at the Highway Policy Zone level, presenting EC, SC, GEC, and SERR equal 0.145, -0.018, 0.013, and 0.117, respectively.

The evaluation carried out so far suggests that PAC highway investments generated solid and positive effects on the economy, even considering social and environmental issues and several institutional and political constraints that have hampered the program's implementation and effectiveness (Amann *et al.*, 2016; Armijo and Rhodes, 2017; Burrier, 2019; Raiser *et al.*, 2017). On the other hand, the PAC social component is negative, and the environmental component is higher than the simulations' averages regardless of the regional scale used.

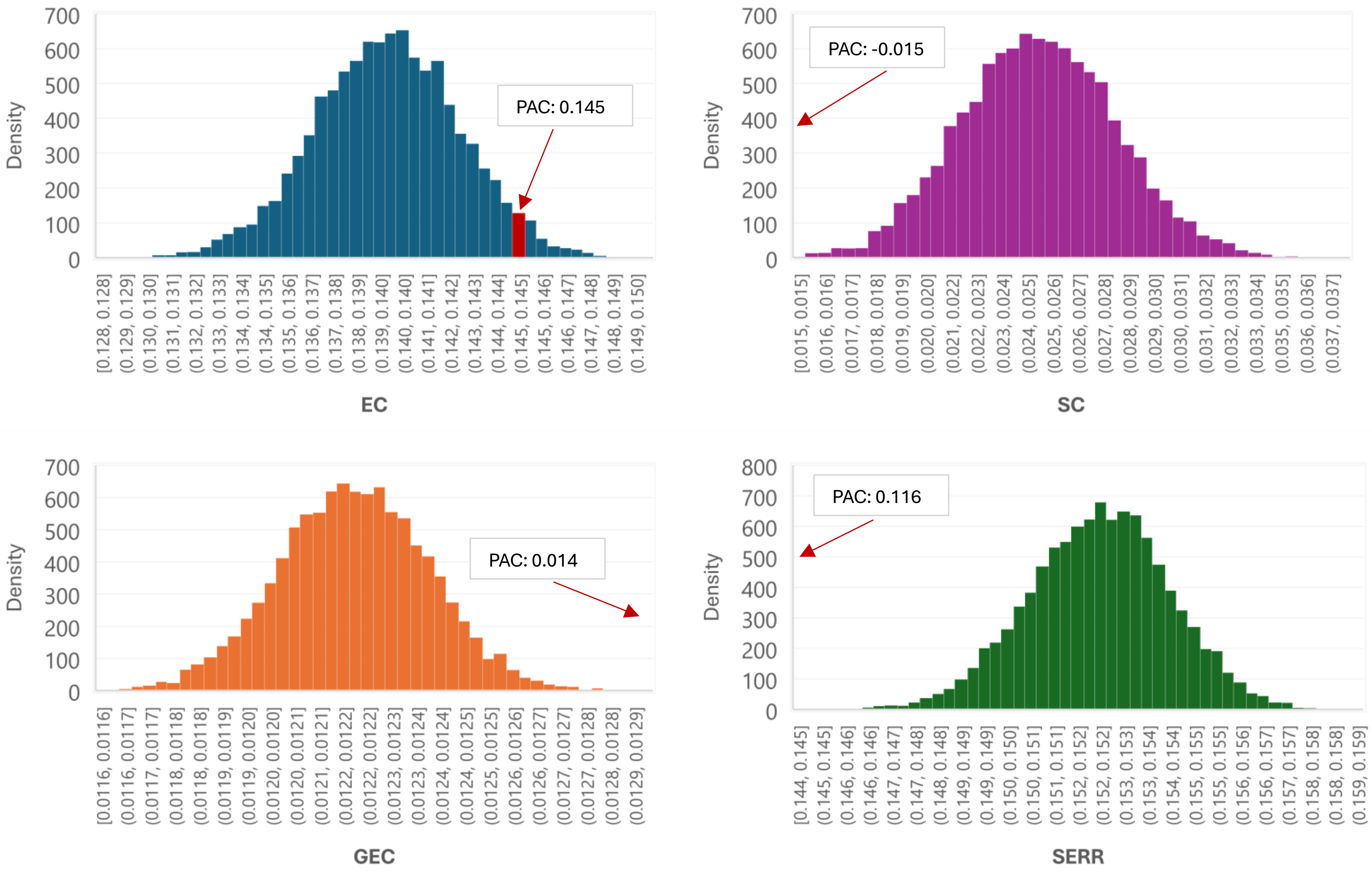
In this sense, we ask whether the PAC could have achieved better outcomes through a more redistributive, equitable, and environmentally friendly road allocation. To answer this question, we reassess the simulation findings for eligible RGIs (Figure 5.7) and calculate the differences between the PAC's economic, social, and environmental components and the simulation averages. We consider the average values of the simulation for eligible RGIs as benchmarks. Figure 5.8 shows the results.

Figure 5.6. Simulation results: EC, SC, GEC and SERR



Source: authors' elaboration.

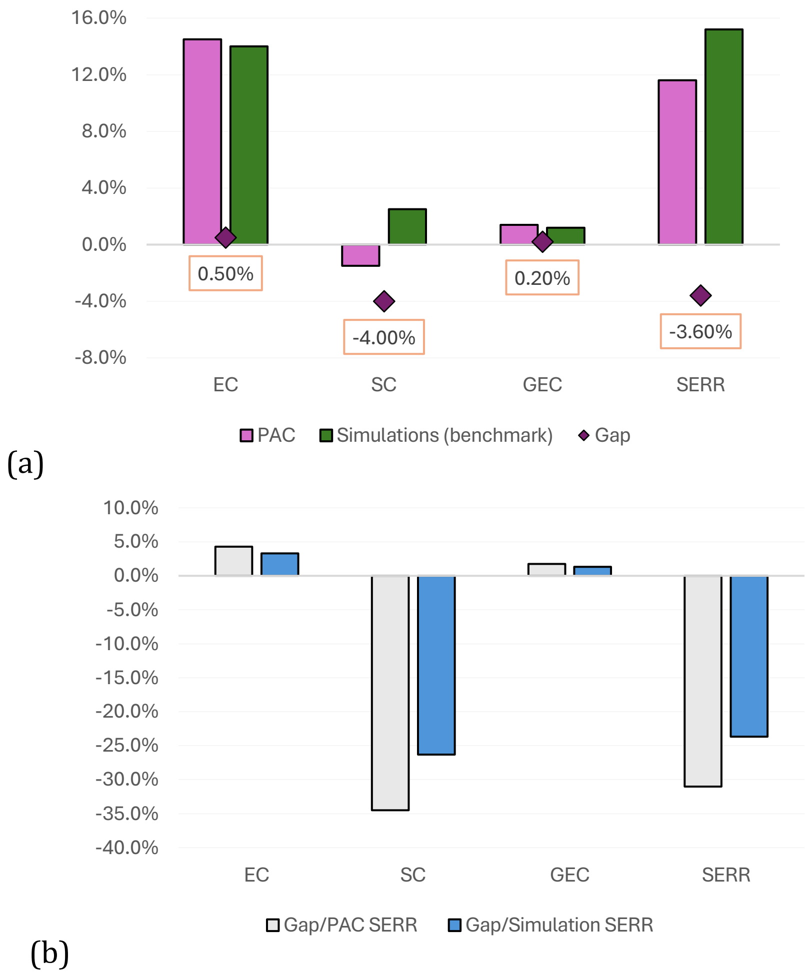
Figure 5.7. Simulation results for eligible RGIs: EC, SC, GEC, and SERR



Source: authors' elaboration.

In Figure 5.8a, the symbol in orange describes the gap between the PAC and simulation results. In economic terms, results suggest a slightly positive difference of 0.5 percentage points in favor of the PAC due to its economic bias. On the other hand, the social component presents a substantial gap of 4 percentage points, while the environmental cost gap is 0.20, reaffirming that the PAC could have reached better social and environmental results by focusing on priority localities. The overall result is a gap of 3.6 percentage points in the PAC SERR in relation to the prioritized regions simulations SERR.

Figure 5.8. The PAC gains (losses): differences (gap) between PAC and average simulations for eligible RGIs SERRs (a) and ratios between the gap and the SERR (b)



Source: authors' elaboration.

Figure 5.8b exhibits the ratios between the gap values and the PAC and simulation SERRs. These measures allow us to identify how much (in percentage terms) of the road investment return has "vanished" by not guiding road allocation on economic, social, and environmental priority areas or how much (in percentage terms) the PAC could have gained by placing its interventions in priority places. The total gap represents 23.68% of the simulation SERR, suggesting that almost one-quarter of the road profitability in a priority allocation scenario would be "wasted" due to an inefficient road placement following the "old" PAC. Similarly, the total gap is 31.03% of the PAC SERR, indicating that the program could have been almost a third more profitable if it had focused on priority regions. It should be noted that those ratios would be even higher if our simulations had considered RGIs classified as priority one only or priorities 1 to 4. The SC is the most critical component in explaining the SERR gap, with a (negative) contribution of 34.48% and 26.32% taking the PAC SERR and Simulation SERR, respectively. The EC represents a (positive) contribution of 4.31% and 3.29%, while the GEC represents a (negative) contribution of 1.72% and 1.32%. Then, this exercise corroborates previous findings indicating the inadequate PAC investments prioritization in ineffectively attending poorer and less infrastructure-endowed regions.

In brief, the *ex-post* evaluation conducted in this section allows us to state that the PAC was an economically effective program, investing in regions with returns significantly above the conventionally used cut-off return rates. Nonetheless, those economic benefits are, to some degree, counterbalanced by harmful social and environmental components. Then, higher road investment returns could have been achieved by allocating investments to priority regions, ensuring a *win-win* scenario marked by a more equitable and environmentally friendly economic growth.

5.6. Concluding remarks

We proposed an empirical strategy to rank priority regions for road investments in Brazil. Unlike previous papers focusing on economic issues, we considered those places where inclusive and sustainable economic growth might arise from road investments as priorities. Our analysis is conducted at different spatial scales, providing insights into local, regional, and national road policies. Due to the pronounced infrastructure deficit in Brazil, we identified eligible regions for road investments in different parts of the country. The economic component is generally greater in the South and Southeast, the social component is larger in the North and Northeast, and the environmental costs are superior in the North and Center-West. By evaluating the returns on investment at a more local level (RGIs), it is possible to identify regions where road investments can generate growth, reduce inequalities, and respect the environment. When the analysis is more aggregated at the state or zone level, it becomes more challenging to find *win-win* combinations, with regions showing at least one weakness in terms of the three components evaluated. This result sheds some light on the huge complexity of implementing national infrastructure policies aimed at building, improving, or granting to private partners extensive road segments, which will cross different regions of the country with diverse economic, social, and environmental attributes.

Next, we conducted an *ex-post* evaluation of PAC 1 and 2 (2007-2018), applying our prioritization criteria. From this, we clarified some ways to use our results for public policy aims. Our PAC evaluation suggests that the program was economically effective, providing economic returns above the commonly used cut-off return rates. However, the economic benefits were somewhat offset by constrained social and environmental components, suggesting that the PAC could have achieved equal or better results by focusing on poorer regions with less infrastructure stock. We calculated a gap in the PAC return rate compared with a hypothetical priority region return rate equal to 3.6%, suggesting that the program profitability could have been increased by 30% by allocating investment to priority RGIs in economic, social, and environmental terms.

While we contribute to the empirical literature on infrastructure and regional development in some ways, gaps remain. We provided several inputs for transportation policy planning, designing, funding, and evaluation. From this, future research might focus on the best road management model (public, private, PPP, and so forth), depending on the economic, social, and environmental components. In addition, succeeding studies can use our inputs to explore cross-subsidies viability in the Brazilian road transportation sector, identifying the road investment rentability by region or road segment. Finally, our analysis can be extended to evaluate new road programs as the New PAC.

6. CONCLUSION

This dissertation aims to investigate the role of transportation infrastructure investments in promoting inclusive and sustainable economic growth. We construct a novel dataset at the municipal level and employ a three-step IV identification strategy to infer causal road impacts on economic activity, regional inequality, and the environment. From those empirical inputs, we calculate regionalized sustainable and equitable return rates to highway investments in Brazil.

We have some broad results. First, road infrastructure development successfully increases productivity. However, this impact is likely biased when not considering measurement error in infrastructure variables and the non-random placement nature of transportation policies. Second, road impacts on productivity are larger for poorer and less infrastructure-endowed regions. This finding points out that road interventions can be used to promote economic growth and reduce regional disparities. Third, the road impacts on the environment proved to be harmful. The damaging effect of roads on GHG emissions is higher for poorer and more remote regions, suggesting the existence of some trade-off between social bonuses by reducing inequalities and environmental costs by raising pollution. Fourth, we compute novel sustainable and equitable return rates to highway investments and provide criteria for classifying Brazilian regions in terms of maximizing (minimizing) economic, social, and environmental benefits (costs). From these return rates, we identify *win-win* localities wherein economic growth is expected to be achieved with reduced regional inequalities and minimized environmental damages.

These findings pose an essential question: can road infrastructure policies promote inclusive and sustainable economic growth? This is particularly important for developing countries that are marked by deep regional unevenness and environmental issues, such as Brazil. If public authorities aim to increase welfare, reduce inequality, and avoid environmental degradation, our results point out some best-case scenarios wherein those outcomes can work together. Nonetheless, in most cases, governments will have to deal with economic, social, or environmental issues undermining road investment returns. Then, establishing priority areas plays a pivotal role in maximizing the broader returns to road investments. This dissertation develops new tools to guide policymakers in planning, designing, financing, executing, and evaluating more efficient, inclusive, and sustainable transportation policies. Empirical exercises in all chapters can be replicated for other developing countries upon data availability and adapted for other Brazilian infrastructure sectors and subsectors such as sanitation, power, telecommunications, railroads, airports, and ports.

I. REFERENCES

Adler, M. W., Pasidis, I., Levkovich, O., Lembcke, A. C., & Ahrend, R. (2020). Roads, market access and regional economic development.

Alam, M., Dappe, M. H., Melecky, M., & Goldblatt, R. (2022). Wider economic benefits of transport corridors: Evidence from international development organizations. *Journal of Development Economics*, 158, 102900.

Albalate, D., & Fageda, X. (2016). High-technology employment and transportation: Evidence from the European regions. *Regional studies*, 50(9), 1564-1578.

Allen, T., & Arkolakis, C. (2014). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics*, 129(3), 1085-1140.

Amann, E., Baer, W., Trebat, T., & Lora, J. V. (2016). Infrastructure and its role in Brazil's development process. *The Quarterly Review of Economics and Finance*, 62, 66-73.

Anas, A. (2020). The cost of congestion and the benefits of congestion pricing: A general equilibrium analysis. Transportation Research Part B: Methodological, 136, 110-137.

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of mathematical statistics*, 20(1), 46-63.

Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11, 727-753.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic perspectives*, 15(4), 69-85.

Arbués, P., Baños, J. F., & Mayor, M. (2015). The spatial productivity of transportation infrastructure. *Transportation Research Part A: Policy and Practice*, 75, 166-177.

Arestis, P., Ferrari-Filho, F., Resende, M. F. D. C., & Bittes Terra, F. H. (2022). A critical analysis of the Brazilian 'expansionary fiscal austerity': why did it fail to ensure economic growth and structural development?. *International Review of Applied Economics*, 36(1), 4-16.

Armijo, L. E., & Rhodes, S. D. (2017). Explaining infrastructure underperformance in Brazil: Cash, political institutions, corruption, and policy Gestalts. *Policy Studies*, 38(3), 231-247.

Aschauer, D. A. (1989). Does public capital crowd out private capital? *Journal of monetary economics*, 24(2), 171-188.

Asher, S., Garg, T., & Novosad, P. (2020). The ecological impact of transportation infrastructure. *The Economic Journal*, 130(629), 1173-1199.

Asher, S., & Novosad, P. (2020). Rural roads and local economic development. *American Economic Review*, 110(3), 797-823.

Assunção, R. M., Neves, M. C., Câmara, G., & da Costa Freitas, C. (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, 20(7), 797-811.

Awad, A., Mallek, R. S., Ozturk, I., & Abdalla, Y. A. (2023). Infrastructure Development's role in environmental degradation in sub-Saharan Africa: Impacts and transmission channels. *Journal of Cleaner Production*, 137622.

Azevedo, T. R., Costa Junior, C., Brandão Junior, A., Cremer, M. D. S., Piatto, M., Tsai, D. S., ... & Kishinami, R. (2018). SEEG initiative estimates of Brazilian greenhouse gas emissions from 1970 to 2015. *Scientific data*, 5(1), 1-43.

Banerjee, A., Duflo, E., & Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics*, 145, 102442.

Barzin, S., D'Costa, S., & Graham, D. J. (2018). A pseudo–panel approach to estimating dynamic effects of road infrastructure on firm performance in a developing country context. *Regional Science and Urban Economics*, 70, 20-34.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2), 775-805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, 99(3), 435-448.

Baum-Snow, N., Henderson, J. V., Turner, M. A., Zhang, Q., & Brandt, L. (2020). Does investment in national highways help or hurt hinterland city growth? *Journal of Urban Economics*, 115, 103124.

Behrens, K., Lamorgese, A., Ottaviano, G., & Tabuchi, T. (2007). "Testing the 'home market effect' in a multi-country world". CORE-UCL, Louvain, mimeo.

Behrens, K., Lamorgese, A. R., Ottaviano, G. I., & Tabuchi, T. (2007). Changes in transport and non-transport costs: Local vs global impacts in a spatial network. *Regional Science and Urban Economics*, 37(6), 625-648.

Bird, J., & Straub, S. (2020). The Brasilia experiment: the heterogeneous impact of road access on spatial development in Brazil. *World Development*, 127, 104739.

Brasil. Exposição geral da situação econômica do Brasil – 1957. Rio de Janeiro: Conselho Nacional de Economia, 1958.

Brasil. Política de transportes e comunicações do governo do presidente Juscelino Kubitschek de Oliveira no quinquênio 1956-60. Rio de Janeiro: Departamento de Imprensa Nacional/MVOP, 1956a.

Brasil. Plano Quinquenal de Obras Rodoviárias Federais – 1956/60. Rio de Janeiro: MVOP/DNER, 1956b.

Brooks, L., & Liscow, Z. (2019). Is Infrastructure Spending Like Other Spending? In: *Economics of Infrastructure Investment*. University of Chicago Press.

Burrier, G. (2019). Politics or technical criteria? The determinants of infrastructure investments in Brazil. *The Journal of Development Studies*, 55(7), 1436-1454.

Cadot, O., Röller, L. H., & Stephan, A. (2006). Contribution to productivity or pork barrel? The two faces of infrastructure investment. *Journal of Public Economics*, 90(6-7), 1133-1153.

Calderón, C., & Servén, L. (2014). Infrastructure, growth, and inequality: An overview. World Bank Policy Research Working Paper, (7034).

Carranza, L., Daude, C., & Melguizo, A. (2014). Public infrastructure investment and fiscal sustainability in Latin America: incompatible goals? *Journal of Economic Studies*, 41(1), 29-50.

Castells, A., & Solé-Ollé, A. (2005). The regional allocation of infrastructure investment: The role of equity, efficiency, and political factors. *European Economic Review*, 49(5), 1165-1205.

Chandra, A., & Thompson, E. (2000). Does public infrastructure affect economic activity? Evidence from the rural interstate highway system. *Regional science and urban economics*, 30(4), 457-490.

Churchill, S. A., Inekwe, J., Ivanovski, K., & Smyth, R. (2021). Transport infrastructure and CO2 emissions in the OECD over the long run. *Transportation Research Part D: Transport and Environment*, 95, 102857.

Coşar, A. K., Demir, B., Ghose, D., & Young, N. (2022). Road capacity, domestic trade, and regional outcomes. *Journal of Economic Geography*, 22(5), 901-929.

Cosci, S., & Mirra, L. (2018). A spatial analysis of growth and convergence in Italian provinces: the role of road infrastructure. *Regional Studies*, 52(4), 516-527.

da Silva, C. F. A., de Andrade, M. O., dos Santos, A. M., & de Melo, S. N. (2023). Road network and deforestation of indigenous lands in the Brazilian Amazon. *Transportation Research Part D: Transport and Environment*, 119, 103735.

Dijkstra, E. W. (2022). A note on two problems in connexion with graphs. In Edsger Wybe Dijkstra: His Life, Work, and Legacy (pp. 287-290).

Donaldson, D., & Hornbeck, R. (2016). Railroads and American economic growth: A “market access” approach. *The Quarterly Journal of Economics*, 131(2), 799-858.

Duranton, G., & Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, 79(4), 1407-1440.

Duranton, G., Morrow, P. M., & Turner, M. A. (2014). Roads and Trade: Evidence from the US. *Review of Economic Studies*, 81(2), 681-724.

Emodi, N. V., Inekwe, J. N., & Zakari, A. (2022). Transport infrastructure, CO2 emissions, mortality, and life expectancy in the Global South. *Transport Policy*, 128, 243-253.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. *Review of Economic Studies*, 81(3), 1046-1070.

Fageda, X., & Olivieri, C. (2019). Transport infrastructure and regional convergence: A spatial panel data approach. *Papers in Regional Science*, 98(4), 1609-1631.

Farhadi, M. (2015). Transport infrastructure and long-run economic growth in OECD countries. *Transportation Research Part A: Policy and Practice*, 74, 73-90.

Fedderke, J. W., & Bogetić, Ž. (2009). Infrastructure and growth in South Africa: Direct and indirect productivity impacts of 19 infrastructure measures. *World Development*, 37(9), 1522-1539.

Feiveson, A. H. (1999). FAQ: What is the delta method and how is it used to estimate the standard error of a transformed parameter.

Felton, C., & Stewart, B. M. (2022). Handle with Care: A Sociologist's Guide to Causal Inference with Instrumental Variables.

Fernald, J. G. (1999). Roads to prosperity? Assessing the link between public capital and productivity. *American economic review*, 89(3), 619-638.

Ferrante, L., Andrade, M. B., & Fearnside, P. M. (2021). Land grabbing on Brazil's Highway BR-319 as a spearhead for Amazonian deforestation. *Land use policy*, 108, 105559.

Foster, V., Gorgulu, N., Straub, S., & Vagliasindi, M. (2023a). The Impact of Infrastructure on Development Outcomes: A Qualitative Review of Four Decades of Literature.

Foster, V., Gorgulu, N., Straub, S., & Vagliasindi, M. (2023b). The Impact of Infrastructure on Development Outcomes: A Meta-Analysis.

Frischtak, C. R., & Mourão, J. (2017). O Estoque de Capital de Infraestrutura no Brasil: uma abordagem setorial. DESAFIOS DA NAÇÃO, 133.

Frye, D. (2016). Transportation networks and the geographic concentration of employment.

Garcia-López, M. À., Holl, A., & Viladecans-Marsal, E. (2015). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics*, 85, 52-67.

Georgatzi, V. V., Stamboulis, Y., & Vetsikas, A. (2020). Examining the determinants of CO2 emissions caused by the transport sector: Empirical evidence from 12 European countries. *Economic Analysis and Policy*, 65, 11-20.

Gertler, P. J., Gonzalez-Navarro, M., Gracner, T., & Rothenberg, A. D. (2019). Road Quality, Local Economic Activity, and Welfare: Evidence from Indonesia's Highways.

Gertler, P., Gonzalez-Navarro, M., Gracner, T., & Rothenberg, A. D. (2022). Road Maintenance and Local Economic Development: Evidence from Indonesia’s Highways (No. w30454). National Bureau of Economic Research.

Ghani, E., Goswami, A. G., & Kerr, W. R. (2016). Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing. *The Economic Journal*, 126(591), 317-357.

Ghannouchi, I., Ouni, F., & Aloulou, F. (2023). Investigating the impact of transportation system and economic growth on carbon emissions: application of GMM system for 33 European countries. *Environmental Science and Pollution Research*, 30(39), 90656-90674.

Golden, M. A., & Picci, L. (2008). Pork-barrel politics in postwar Italy, 1953–94. American *Journal of Political Science*, 52(2), 268-289.

Greene, D.L., Park, S., Liu, C., 2014. Analyzing the transition to electric drive vehicles in the U.S. Futures 58, 34–52.

Han, J., Meng, X., Liu, J., & Zhang, Y. (2017). The impact of infrastructure stock density on CO2 emissions: Evidence from China provinces. *Sustainability*, 9(12), 2312.

Herzog, I. (2021). National transportation networks, market access, and regional economic growth. *Journal of Urban Economics*, 122, 103316.

Hirschman, A. (1958). O. “The Strategy of Economic Development”. Yale University Press.

Holl, A. (2012). Market potential and firm-level productivity in Spain. *Journal of Economic Geography*, 12(6), 1191-1215.

Holl, A. (2016). Highways and productivity in manufacturing firms. *Journal of Urban Economics*, 93, 131-151.

Hsu, W. T., & Zhang, H. (2014). The fundamental law of highway congestion revisited: Evidence from national expressways in Japan. *Journal of urban economics*, 81, 65-76.

Huang, Y., & Xiong, W. (2017). Geographic Distribution of Firm Productivity and Production: A ", Market Access" Approach. Working Paper.

Iasco Pereira, H. C., Romero, J. P., & Medeiros, V. (2021). Kaldor–Verdoorn's law and institutions: evidence from Brazilian municipalities. *Cambridge Journal of Economics*, 45(3), 511-536.

Imbens, G. (2014). Instrumental variables: an econometrician's perspective (No. w19983). National Bureau of Economic Research.

Jaworski, T., & Kitchens, C. T. (2019). National policy for regional development: Historical evidence from Appalachian highways. *Review of Economics and Statistics*, 101(5), 777-790.

Ke, X., & Yan, Y. (2021). Can proactive fiscal policy achieve the goal of "Beyond Keynesianism"? Evidence from China's National Highway System. *Review of Development Economics*, 25(2), 1078-1103.

Kemmerling, A., & Stephan, A. (2002). The contribution of local public infrastructure to private productivity and its political economy: Evidence from a panel of large German cities. *Public Choice*, 113(3-4), 403-424.

Kenny, C. (2009). Measuring corruption in infrastructure: Evidence from transition and developing countries. *The Journal of Development Studies*, 45(3), 314-332.

Kornfeld, R., & Fraumeni, B. M. (2022). How Should We Measure Infrastructure? The Case of Highways and Streets (No. w30045). National Bureau of Economic Research.

Krugman, P. (1993). First nature, second nature, and metropolitan location. *Journal of regional science*, 33(2), 129-144.

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical society*, 7(1), 48-50.

Laird, J. J., & Venables, A. J. (2017). Transport investment and economic performance: A framework for project appraisal. *Transport policy*, 56, 1-11.

Lee, J. K. (2021). Transport infrastructure investment, accessibility change and firm productivity: Evidence from the Seoul region. *Journal of Transport Geography*, 96, 103182.

Li, H., & Luo, N. (2022). Will improvements in transportation infrastructure help reduce urban carbon emissions?——motor vehicles as transmission channels. *Environmental Science and Pollution Research*, 29(25), 38175-38185.

Li, Z., Wu, M., & Chen, B. R. (2017). Is road infrastructure investment in China excessive? Evidence from productivity of firms. *Regional Science and Urban Economics*, 65, 116-126.

Lu, H., Zhao, P., Hu, H., Zeng, L., Wu, K. S., & Lv, D. (2022). Transport infrastructure and urban-rural income disparity: A municipal-level analysis in China. *Journal of transport geography*, 99, 103292.

Lima, M., Santana, D. C., Junior, I. C. M., Costa, P. M. C. D., Oliveira, P. P. G. D., Azevedo, R. P. D., ... & Silva Junior, C. A. D. (2022). The "New Transamazonian Highway": BR-319 and Its Current Environmental Degradation. *Sustainability*, 14(2), 823.

Lin, B., & Chen, Y. (2020). Transportation infrastructure and efficient energy services: A perspective of China's manufacturing industry. *Energy Economics*, 89, 104809.

Lin, C., Liu, G., & Müller, D. B. (2017). Characterizing the role of built environment stocks in human development and emission growth. *Resources, Conservation and Recycling*, 123, 67-72.

Luo, Z., Wan, G., Wang, C., & Zhang, X. (2018). Urban pollution and road infrastructure: A case study of China. *China Economic Review*, 49, 171-183.

Martín-Barroso, D., Núñez-Serrano, J. A., & Velázquez, F. J. (2015). The effect of accessibility on productivity in Spanish manufacturing firms. *Journal of Regional Science*, 55(5), 708-735.

Martincus, C. V., Carballo, J., & Cusolito, A. (2017). Roads, exports and employment: Evidence from a developing country. *Journal of Development Economics*, 125, 21-39.

Medeiros, V., & Ribeiro, R. S. M. (2020). Power infrastructure and income inequality: Evidence from Brazilian state-level data using dynamic panel data models. *Energy Policy*, 146, 111734.

Medeiros, V., Ribeiro, R. S. M., & do Amaral, P. V. M. (2021a). Infrastructure and household poverty in Brazil: A regional approach using multilevel models. *World Development*, 137, 105118.

Medeiros, V., Ribeiro, R. S., & AMARAL, P. (2021b). Productivity of transportation infrastructure in Brazil: a sectoral and regional approach using dynamic panel data models. Coletânea de Trabalhos–Prêmio ABDE-BID: edição.

Medeiros, V., de Queiroz Stein, A., Ribeiro, R. S. M., & Dos Santos, U. P. (2021c). Transportation infrastructure and University-Industry collaborations: regional-level evidence from Brazil. Blucher Engineering Proceedings.

Medeiros, V., Saulo Marques Ribeiro, R., & Vasconscelos Maia do Amaral, P. (2022). Infrastructure and income inequality: An application to the Brazilian case using hierarchical spatial autoregressive models. *Journal of Regional Science*, 62(5), 1467-1486.

Medeiros, V; Ribeiro, R. S. M.; Amaral, P. V. M.; Stein, A. Q. (2024). Highway Infrastructure and Local Outcomes: measuring causal impacts of infrastructure investments using a three-step instrumental variable identification strategy. São Paulo: Centro de Pesquisa em Macroeconomia das Desigualdades (Made/USP), 2024. (Working Paper, n.23).

Medeiros, V; Ribeiro, R. S. M.; Amaral, P. V. M.; Stein, A. Q. (*forthcoming*). Infrastructure, growth, and regional disparities: looking at the efficiency, redistribution, and equity goals in road Investments. São Paulo: Centro de Pesquisa em Macroeconomia das Desigualdades (Made/USP). (Working Paper, s.n.).

Meijers, E., Hoekstra, J., Leijten, M., Louw, E., & Spaans, M. (2012). Connecting the periphery: Distributive effects of new infrastructure. *Journal of Transport*

Geography, 22, 187-198.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, 90(4), 683-701.

Ministry of Economy. 2021. Social discount rate for evaluating infrastructure investments: update after public consultation. consultation. SEI Technical Note no. 19911/2020/ME.

Ministry of Economy (2022). General guide to socio-economic cost-benefit analysis of infrastructure investment projects. Brasília : SDI/ME.

Monastiriotis, V., & Psycharis, Y. (2014). Between equity, efficiency and redistribution: An analysis of revealed allocation criteria of regional public investment in Greece. *European Urban and Regional Studies*, 21(4), 445-462.

Nassif, A., Feijó, C., & Araújo, E. (2015). Structural change and economic development: is Brazil catching up or falling behind? *Cambridge Journal of Economics*, 39(5), 1307-1332.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. Proceedings of the National Academy of Sciences, 114(7), 1518-1523.

Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3), 358-369.

Ottaviano, G. I. (2008). Infrastructure and economic geography: An overview of theory and evidence. EIB papers, 13(2), 8-35.

Percoco, M. (2016). Highways, local economic structure, and urban development. *Journal of Economic Geography*, 16(5), 1035-1054.

Pizer, S. D. (2016). Falsification testing of instrumental variables methods for comparative effectiveness research. *Health services research*, 51(2), 790-811.

Puga, D. (1999). "The rise and fall of regional inequalities". *European Economic Review*, (43:2), pp. 303-334.

Quadros, S. G. R., & Nassi, C. D. (2015). An evaluation on the criteria to prioritize transportation infrastructure investments in Brazil. *Transport Policy*, 40, 8-16.

Raiser, M., Clarke, R., Procee, P., Briceno-Garmendia, C., Kikoni, E., Kizito, J., & Viñuela, L. (2017). Back to Planning.

Redding, S. J., & Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9, 21-58.

Redding, S. J., & Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. *Handbook of regional and urban economics*, 5, 1339-1398.

Rennert, K., Errickson, F., Prest, B. C., Rennels, L., Newell, R. G., Pizer, W., ... & Anthoff, D. (2022). Comprehensive evidence implies a higher social cost of CO2. *Nature*, 610(7933), 687-692.

Ricke, K., Drouet, L., Caldeira, K., & Tavoni, M. (2018). Country-level social cost of carbon. *Nature Climate Change*, 8(10), 895-900.

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of economic growth*, 9, 131-165.

Rokicki, B., & Stępniak, M. (2018). Major transport infrastructure investment and regional economic development–An accessibility-based approach. *Journal of Transport Geography*, 72, 36-49.

Roberts, M., Melecky, M., Bougna, T., & Xu, Y. (2020). Transport corridors and their wider economic benefits: A quantitative review of the literature. *Journal of Regional Science*, 60(2), 207-248.

Santos, G. (2017). Road transport and CO2 emissions: What are the challenges?. *Transport Policy*, 59, 71-74.

Sharif, F., & Tauqir, A. (2021). The effects of infrastructure development and carbon emissions on economic growth. *Environmental Science and Pollution Research*, 28, 36259-36273.

Sheard, N. (2014). Airports and urban sectoral employment. *Journal of Urban Economics*, 80, 133-152.

Shi, H., & Huang, S. (2014). How much infrastructure is too much? A new approach and evidence from China. *World Development*, 56, 272-286.

Souza Jr, C. M., Z. Shimbo, J., Rosa, M. R., Parente, L. L., A. Alencar, A., Rudorff, B. F., ... & Azevedo, T. (2020). Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. *Remote Sensing*, 12(17), 2735.

Storeygard, A. (2016). Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa. *The Review of economic studies*, 83(3), 1263-1295.

Straub, S. (2008). Infrastructure and growth in developing countries: recent advances and research challenges. World Bank policy research working paper, (4460).

Straub, S. (2011). Infrastructure and development: A critical appraisal of the macro-level literature. *The Journal of Development Studies*, 47(5), 683-708.

UNEP (2014) Valuing Plastics: The Business Case for Measuring, Managing and Disclosing Plastic Use in the Consumer Goods Industry.

Wang, Z., Wu, G. L., & Feng, Q. (2020). Productivity of core infrastructure investment in China: An input–output approach. *The World Economy*, 43(12), 3384-3406.

Welde, M., & Tveter, E. (2022). The wider local impacts of new roads: A case study of 10 projects. *Transport policy*, 115, 164-180.

Xiao, F., Pang, Z., Yan, D., Kong, Y., & Yang, F. (2023). How does transportation infrastructure affect urban carbon emissions? an empirical study based on 286 cities in China. *Environmental Science and Pollution Research*, 30(4), 10624-10642.

Xie, R., Fang, J., & Liu, C. (2017). The effects of transportation infrastructure on urban carbon emissions. *Applied Energy*, 196, 199-207.

Xu, H., Cao, S., & Xu, X. (2022). The development of highway infrastructure and CO2 emissions: The mediating role of agglomeration. *Journal of Cleaner Production*, 337, 130501.

Xu, M. I. N. G., & Feng, Y. U. A. N. (2022). How transportation infrastructure affects firm Productivity? Evidence from China. *China Economic Quarterly International*, 2(1), 55-69.

Yamano, N., & Ohkawara, T. (2000). The regional allocation of public investment: Efficiency or equity?. *Journal of Regional Science*, 40(2), 205-229.

Yang, Y. (2018). Transport infrastructure, city productivity growth and sectoral reallocation: Evidence from China. International Monetary Fund.

Yao, S., Zhang, X., & Zheng, W. (2023). On transportation, economic agglomeration, and CO2 emissions in China, 2003–2017. *Environmental Science and Pollution Research*, 30(14), 40987-41001.

Young, A. (2022). Consistency without inference: Instrumental variables in practical application. *European Economic Review*, 147, 104112.

Zhang, X., Hu, Y., & Lin, Y. (2020). The influence of highway on local economy: Evidence from China's Yangtze River Delta region. *Journal of Transport Geography*, 82, 102600.

I. APPENDIX - HIGHWAY INFRASTRUCTURE AND ECONOMIC DEVELOPMENT

A. Study Design (the first step)

Our main goal is to estimate causal impacts of highway investments on local outcomes. To this end, we rely on a national road program (a program within the PAC) implemented by the Brazilian Federal Government between 2007 and 2018.

By desegregating national investments at the municipal level, procedure we describe in detail further ahead in the data section, we can evaluate local impacts of an extensive and aggregated road policy. This study design is important to alleviate endogeneity concerns between infrastructure investments and economic activity (Faber, 2014; Herzog, 2021), as we do not expect municipalities directly influencing the Brazilian federal government decisions at placing highways across the country.

Currently, Brazil has 5,570 municipalities into 27 states. Whilst we might expect some important and direct political influence of the state level governments on the federal government investment priorities potentially biasing econometric estimates, this issue is quite reduced at using municipal level data.

One evident way that municipalities might influence federal government decisions is by their population and economic activity. Central cities may receive priority treatment from federal infrastructure policies in order to attend massive populations and foster national economic growth. However, the most part of Brazilian municipalities are medium or small sized. In 2006, around 4,986 (89,5%) out of 5,570 municipalities had less than 50,000 inhabitants. Only 130 (2,3%) municipalities had population above 200,000. At average, a municipality had 0,018% of the national population, indicating that municipal population is unlikely to affect federal public policy decisions directly in the vast majority of cases.

Second, elected federal deputies may request policies more directly from the federal government. In this regard, the representatives may act in favor of the municipalities that gave them the most votes. Each Brazilian State elects a certain number of federal deputies by nominal votes. The deputies may demand financial resources for projects in specific regions and cities, as well as influence legislations and norms that will impact infrastructure policies. In this sense, if we observe a high share of votes for a federal deputy coming from a specific municipality, we can suppose that deputy devoting greater efforts to serve his electorate. Nonetheless, this argument seems to hold for a few large municipalities, especially the 27 state capitals. For instance, in the 2006 election in São Paulo – the most populated Brazilian state –, an elected federal deputy received an average of 162,827 nominal votes. Only 17,7% of the municipalities of São Paulo had an electorate greater than that in 2006, indicating that in most cases a reasonable number of municipalities is needed to elect a candidate. In addition, 434 federal deputies were nationally elected, a considerable number that raises doubts on the political strength of a single deputy to directly influence federal decisions in favor of a specific city.

In this sense, evaluating the federal road policy impact on local outcomes strongly alleviates broad endogeneity issues. First, it is unlikely that small and medium sized municipalities have any direct political power to influence federal policies in their favor. Second, even whether we believe a few large and potentially targeted cities influencing directly federal decisions, we can use the inconsequential unit approach to exclude them, which allows us to keep a large number of "non-targeted" observations.

B. Cost-related IV rationality: avoiding (or alleviating) measurement bias of monetary highway investments variables

A way of dealing with measurement error (and maybe omitted variable) bias of road measures is choosing reliable cost-related IVs. Conditional on controls, cost-related IVs – for instance, geographical, environmental, and physical human costs as expropriation and interferences – might affect the outcome variable only through the highway variable (Holl, 2012; Lu *et al.*, 2022; Martín-Barroso, Nunez-Serrano, and Velazquez, 2015; Zhang, Hu, and Lin, 2020). We may expect local terrain ruggedness making a road project unfeasible, which in turn is expected to affect a region or city economic development. Nonetheless, it is unlikely this observable characteristic will directly impact GDP, population, or another outcome variable growth. Similarly, environmental costs – as the extent or existence of legally protected areas – and physical human costs – as expropriation and interferences due to highly urban density – directly affect the feasibility and success of an infrastructure project but are unexpected to directly affect outcome variables.

This kind of IV may avoid (or alleviate) endogeneity bias in two ways. First, they may solve endogeneity issues related to omitted variables bias commonly found in infrastructure-economic development studies, as proposed by past studies (Holl, 2012; Lu *et al.*, 2022; Martín-Barroso, Nunez-Serrano, and Velazquez, 2015; Zhang, Hu, and Lin, 2020). Second, and which we consider to be more reasonable, they might act as a corrective instrument for measurement error bias of highway investment variables. This is particularly relevant when using monetary highway variables as investment flows in developing economies wherein a high inefficiency in allocating infrastructure investments is expected (Calderón and Servén, 2014; Straub, 2011).

Then, instrumentalizing the road variable by the main infrastructure costs may reduce both inefficiency bias - it is expected that more costly locations tend to have greater delays in buildings, making infrastructure investment impacts on outcome variables unclear - and road variable (specially the monetary ones) measurement error bias - for instance, more geographical, environmental and physical human costly locations may obviously demand a higher level of investment *per* length of road, requiring the construction of tunnels, bridges, expropriation and compensation payments, which might be reflected into unclear economic returns.

To better elucidate how cost-related IVs can affect road investments, we use some emblematic Brazilian cases. First, we briefly describe the iconic example of the BR-381/262 highway segment crossing the states of Minas Gerais (MG) and Espírito Santo (ES) to contextualize how cost-related inefficiencies occur in practice. The highway is noted by its poor quality and high traffic accident incidence and death rates, and buildings in many segments of the road has been demanded for decades. According to the CNT 2022 Highway Survey, one of the most critical segments of the highway, between the municipalities of João Monlevade and Martins Soares in the state of Minas Gerais, is in the 345º position in terms of road quality, out of 510 segments evaluated. According to data from the Federal Highway Police (PRF), the BR-381 is the fourth highway in number of accidents and the fifth in accidents with deaths. Due to its critical safety and quality condition, the highway segment is known as the "Highway of Death".

The BR-381/262 (MG/ES) segment is also emblematic because its several risks and project complexity. The road crosses areas with high population and urban infrastructure density, demanding huge additional monetary resources and time efforts to solve expropriation and interferences conflicts. Several delays occurred in the building schedule due to expropriation and evacuation disputes, as many households did not accept the expropriation terms offered by the Brazilian government. Those issues indirectly affected the population welfare through its direct impact on the efficacy and efficiency of the investments in the highway. The conflictive status of the road required a long time to be solved, and supplemental money was needed to both attend the impacted population and maintain unfinished road segments which were not being effectively used by the society.

In 2022, the Brazilian Federal Government tried to grant the BR-381/262 (MG/ES) segment. The result was the auction cancelation due to the lack of private partner interest. It was the fourth time that the “Highway of Death” auction was cancelled. The road sector private agents argued that the project was too risky – as it required massive investments and could suffer from several external interferences –, and it would require higher financial compensation to be taken.

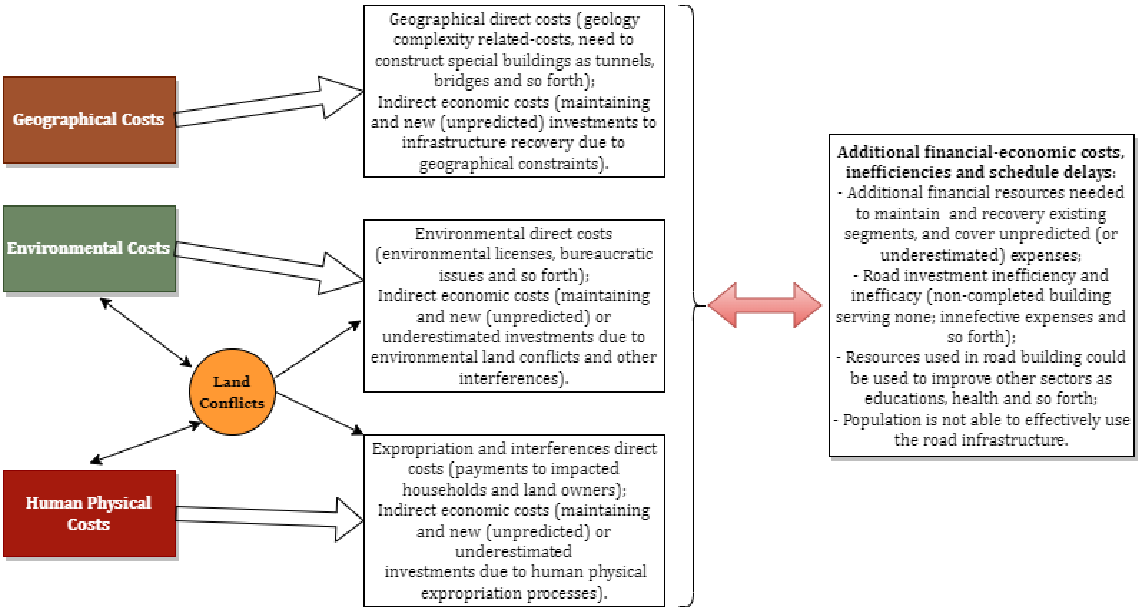
A second interesting case is the BR-163 north segment between the states of Mato Grosso (MT) and Pará (PA). This segment was constructed in the 1970s, but it has never been completely paved until recently. The typical road segment picture was kilometers of stuck trucks for weeks during raining periods. As those trucks mostly carried perishable beans to the Brazilian northern ports, they had several economic losses strictly associated with the poor road condition.

The BR-163 north segment is also symbolic because it crosses a large legal protected indigenous area. To mitigate potential negative environmental and social impacts of the road paving – as deforestation, illegal extractivism and so forth –, the Brazilian government released the “BR-163 Sustainable Plan” (*Plano BR-163 Sustentável*) in 2006. The plan was designed to improve the institutional quality of civil society organizations in the region in terms of its monitoring, evaluation, and information system, as well as expanding the mechanisms of social participation and control.

Due to its critical environmental sensibility, the BR-163 project received massive demands from environmental and indigenous organizations. To cover some of those requirements, the Environment and Renewable Natural Resources (IBAMA) listed 13 conditions to be met by the National Highway Infrastructure Department (DNIT) in paving the road. Among them, it included a study on the need to build passages for animals, a Monitoring Program for Water Courses, an implementation schedule for the Fire Fighting Program, and that the DNIT should inform IBAMA and the National Institute for Colonization and Agrarian Reform (INCRA) on the occurrence of *quilombola* communities found in influence of the highway segment. In practice, those needs implied additional economic and bureaucratic costs. In addition, the buildings in the BR-163 north segment were paralyzed several times due to native people interventions claiming for environmental protection and conservation. It incurred in several schedule delays and conflicts on land rights. Even in 2020, some BR-163 north segments were not fully paved.

Another segment of the BR-163 in the Serra da Caixa Furada in the state of Mato Grosso (MT) presented huge geological issues in a duplication building. In 2014, it occurred a landslide in the road, which implied 5 years of complete stoppage in the building. Due to its geology complexity, the building resumption and the reconstruction of the destructed highway demanded several additional financial efforts. The investments applied in the initial stage were rather inefficient as the population was not able to use the incomplete road duplication. In addition, more money was required to recover the collapsed highway segment and to maintain the already constructed (but not used) road segments. Figure B1 summarizes how geographical, environmental, and physical human costs may impact road infrastructure projects.

Figure B1. Geographical, environmental, and human physical costs in infrastructure projects



Source: authors' elaboration.

The high infrastructure project risk heterogeneity in Brazil, which resulted in empty auctions as in the "Highway of Death" case, forced the Brazilian road sector authorities to rethink the assumptions guiding the rate of return methodology used to establish tariffs and economic-financial balances in the sector. Recent resolutions (6.002/2022, 6.003/2022 and 6.004/2022) on the calculation methodology of the Weighted Average Cost of Capital (WACC) were approved in December 2022 by the National Land Transportation Agency (ANTT).

The pivotal new resolutions aim was to include the main risk components in road infrastructure projects to calculate the federal highway sector WACC. To do so, the ANTT established several measurable indicators related to traffic demand, geographical, environmental, and physical human costs of road projects. For instance, the responsible actors must report the road length crossing legal protected areas, urban areas, and hilly areas. Also, the ANTT 6.002/2022 resolution included measures related to road projects expropriation and interferences costs, illegal urban occupation, and traffic demand risk.

Whilst the new resolutions focus on new road concession projects – including both brownfield and greenfield projects –, they provide us some reliable (and measurable) indicators that can be adapted to predict the feasibility and correct measurement errors of different types of road investments, as building, paving, duplication, enhancements and so forth. As the new ANTT methodology relies on real infrastructure projects, they are based on computable and highly replicable measures. Relying on the infrastructure-development literature (Holl, 2012; Lu *et al.*, 2022; Martín-Barroso, Nunez-Serrano, and Velazquez, 2015; Medeiros *et al.*, 2021b; Zhang, Hu, and Lin, 2020) and measures based on real Brazilian infrastructure project costs, in the next sections we propose some cost-related IVs to overcome endogeneity issues in road investment variables.

In short, we propose several geographical, environmental, and human physical infrastructure project costs to correct measurement error in highway investment variable. As several inefficiencies are likely to occur in the developing country context, we might expect the highway measure to be inflated, implying a higher value *per* kilometer of road. It puzzles the relationship between road infrastructure and economic activity as a larger number of roads (in monetary terms) is needed to generate an additional unit of output, tending to downward biased OLS road elasticity estimates. By instrumentalizing our highway variable by strong and exogenous cost-related instruments, we intend to fix (or alleviate) measurement error bias.

C. Non-random allocation IVs rationality: avoiding (or alleviating) omitted variable bias of highway investment

Even using cost-related IVs to instrumentalize highway investments and potentially correct measurement error, omitted variable bias from non-random road placement may persist. Correcting for measurement error from infrastructure projects inefficiencies is just a (important) part of the problem, but it is likely not enough to eliminate endogeneity bias related to the propensity of certain localities to receive federal highway interventions.

First, governments might allocate roads to underdeveloped regions to promote regionally balanced economic growth or directing highway investments to more developed localities – wherein the expected return to road investment is higher – to foster national economic growth. If it occurs, naïve OLS regressions would be underestimated in the former context, whilst overestimated in the latter.

To correct those issues, several studies have relied on LCP-MST instruments to generate exogenous sources of variation for highway measures based on a global minimization cost network. We can easily apply the LCP-MST method to Brazilian data as well as several other economies. In a more general and replicable way, we can stablish targeted points based on road network data. Starting and ending points of a road receiving federal investments are clearly potential hub candidates. From these core cities, we could calculate LCP-MST hypothetical networks and generate instruments for highway measures. For instance, the capital Brasília is the starting point for all radial roads, whilst for longitudinal, transversal, or diagonal roads, starting and ending points represent economic, touristic, or even political hubs. In this way, this approach might provide a comprehensive and highly replicable way of stablishing hubs in studies whereby targeted cities are not so clear. Other way to identify potential hubs is through historical Census data – for instance, highly populated cities hundreds of years ago – or historical plans. Then, the same LCP-MST procedure might be applied to generate IVs to correct non-random placement of road investments.

In addition, the LCP-MST allow us to include the main infrastructure costs in the minimization cost path. In other words, we might be interested in minimizing the global network cost not only based on Euclidean distances, but also on the main geographical, environmental, and human physical costs. As we argued in the previous section, Brazil has several infrastructure costs and risks constraints, and including those characteristics in the minimization LCP-MST procedure can enhance our constructed instrument and empirical strategy. Some studies have used raster data to include geographical costs in the minimization path (Faber, 2014), approach we improve by adding environmental and human physical indicators.

Second, governments might be guided by political reasons, as to regionally connect the country. If political allocation bias occurs, we could expect OLS estimates be biased towards zero. In Brazil, we have the Brasília Plan case proposed by Bird and Straub (2020) to instrument local road improvements in a political biased road policy context. We also have the Brasília JK Road Cruise, which is an extension of the Brasília Plan and includes some additional targeted cities. In both cases, the rationale behind the instruments is that the national Brazilian government in the 1950s and 1960s aimed to connect the whole country having the new capital Brasília as the central point of the network, and municipalities in the way among Brasília and the end points were incidentally connected. In this sense, we can use those historical plans to generate exogenous instruments – based on the distance from the hypothetical lines connecting cities targeted by the plans to Brasília – for highway investments.

Third, by using highway investment flows disaggregated by intervention types – as building, paving, duplications, enhancements and so forth – as we will propose in this study and describe in further detail in the next section, we also need to correct endogeneity bias from the propensity to a locality already connected by a highway in the start period to receive a road intervention. In the Brazilian case, we might anticipate the federal government prioritizing road segments with critical traffic intensity or traffic accidents. In addition, in the context of underdeveloped economies as Brazil, poor institutions, scarce economic resources, and low planning capacity tends to narrow extensive planned road networks around the country, confining infrastructure interventions to just attend infrastructure demanding localities. If it occurs, municipalities crossed by those critical segments are an obvious intended group, and an upward bias in OLS estimates is expected. We can generate "potential *road intervention areas*" instruments based on traffic safety or traffic intensity data, which we believe are available for several countries. The rationality behind this instrument is that, conditional on controls, municipalities already connected by roads in the start period and farther to "*potential road intervention areas*" are more likely to (inconsequentially) receive highway investments to reduce traffic levels and accidents in the critical areas and its surrounding. However, this "luck" at receiving a federal road intervention would be unrelated to economic or political reasons, providing us a potentially suitable instrument.

In short, we propose a range of LCP-MST, historical and "potential *road intervention areas"* instruments to correct for the non-random placement of roads. As governments might be guided by political and economic reasons, we could observe downward or upward bias in OLS regressions. By using a "free from measurement error" road variable, our proposed non-random allocation instruments have the rough aim to correct for remaining omitted variables biases.

D. National Highway Investments

To measure the impact of national highway investments on local (municipal) economic activity, we construct a new dataset of national investment flows at the municipal level from 2007 to 2018. To this end, we use two main publicly available datasets.

The first one concerns data of investment flows in the highway sector of the Federal Government's Growth Acceleration Program (PAC). This dataset includes annual information on road investment flows for each one of the 27 Brazilian states, including a brief description of each intervention. As an illustrative example, it is reported whether there was a building intervention on Highway 1 between Local X (it can be the name of a municipality, the number of a km of a road or other local as a port, a connection with another road infrastructure as state or municipal and so forth) and Local Y. The data allows us to differentiate interventions in road construction, paving, duplication, enhancements, or maintenance.

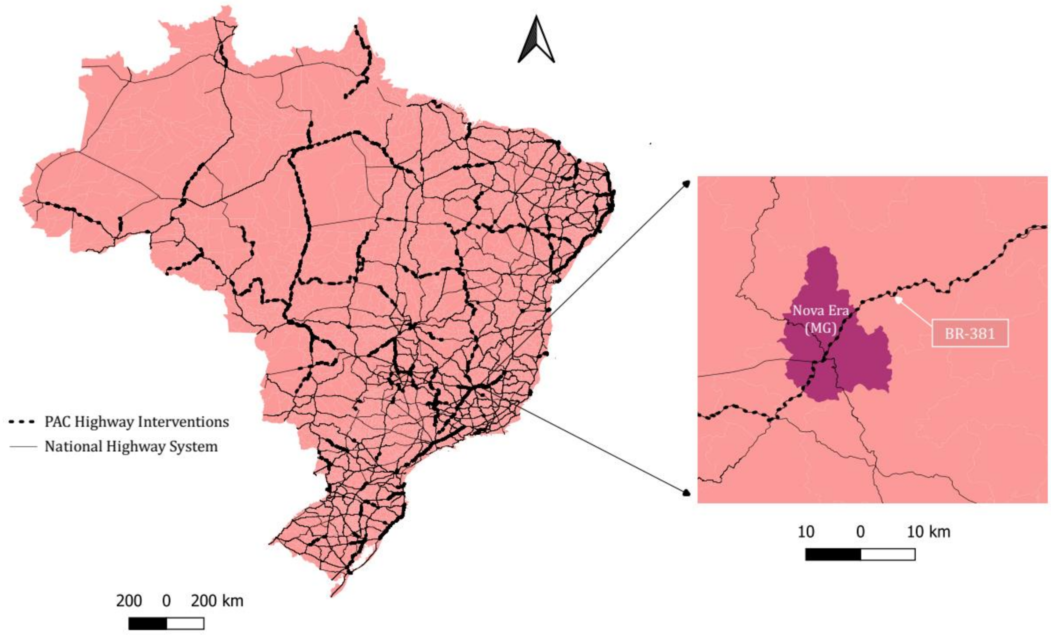
The second one refers to the National Highway System (SNV) georeferenced road data made available by the National Highway Infrastructure Department (DNIT). From this data, it is possible to identify the length of each road segment, its condition (paved, duplicated and so forth), and the places (municipalities) crossed by each one of the intervention road segments.

The PAC dataset is not georeferenced, which implies we have restricted information about whether and in which extent a municipality is crossed by an intervention. To create a national highway investment dataset at the municipal level, we combine the PAC's data description of each intervention and the georeferenced SNV data.

The first step was to identify the treated highway codes and its starting and ending points from the intervention description of the PAC data. Next, the PAC treated highways were linked to the SNV geolocalized data using the highway code and their starting and ending points. It should be noted that the starting and ending points of the two datasets are not fully compatible, which made it necessary to manually match them one by one.

Second, we calculated the total PAC intervention road length by municipality and use it to measure the share of the road length in the municipality in relation to the total intervention road length. As we have investment data only by intervention, we use the measured share to compute the highway investment by municipality. It should be noted that the maintenance intervention descriptions barely describe the state and the highway code. In this sense, it was not possible to geolocate this type of investment at the municipal level, and they were excluded from further analysis. Therefore, the investments refer to building and paving, and duplications and enhancements. The Figure D1 exhibits the PAC highway interventions.

Figure D1. Federal Highway Investments: Municipal Data Construction



Source: authors' elaboration.

To illustrate how the proportion was measured, we show in Figure D1 the example of the municipality of Nova Era, which was crossed by an PAC intervention in the highway BR-381. The total road length covered by one of the buildings in the BR-381 was 163 km, whilst the total investment on it in 2009 was R$ 10.8 million. Around 25 km (15%) of the intervention road segment crosses Nova Era. Multiplying the road intervention proportion of 15% for Nova Era by the total intervention highway investment, the municipality directly received around R$ 1.6 million in investments in 2009. Then, the same procedure was performed annually for all interventions and treated municipalities.

Figure D2 shows the sum of PAC highway investments distributed by municipality between 2007 and 2018. The program directly reached out 703 municipalities, totalizing R$ 77.3 billion, excluding investments in road maintenance. This data is particularly relevant as granular data on infrastructure investment is quite limited (Brooks and Liscow, 2019). In most cases around the world, data is aggregated at the national level, and it is difficult to standardize.

Figure D3 shows the highway investments path during the PAC using the constructed data. The federal highway investments more than duplicated in the PAC's period in comparison with the previous decade (Medeiros *et al.*, 2021b). Our data similarly follows the PAC investments growth pattern[27], presenting a constant rising during the first years. From 2015, the highway PAC investment abruptly dropped following several political and economic Brazilian crisis.

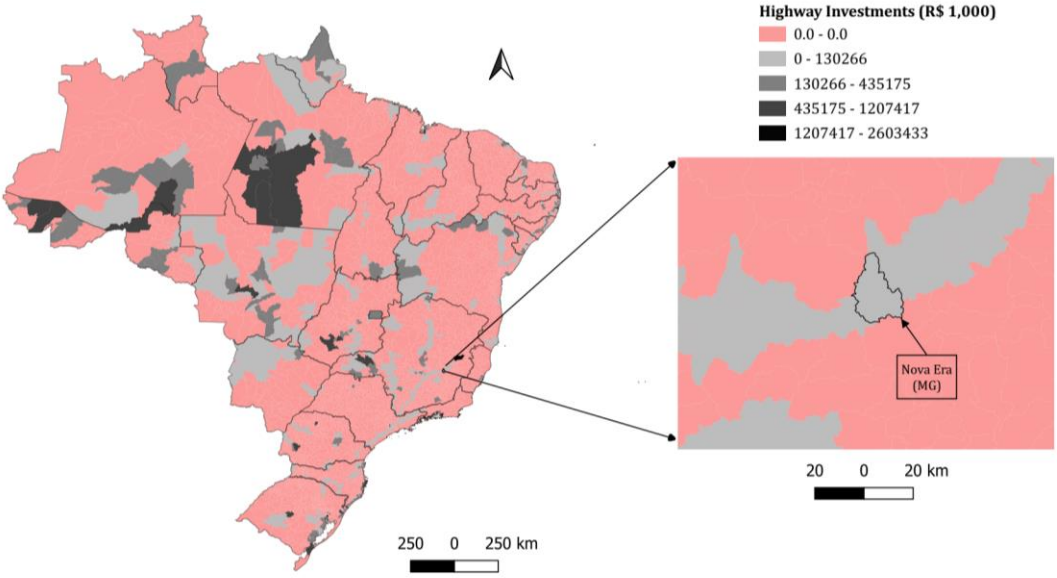
Whilst our municipal data seems to be rather representative and provides a novel and important source to measure local impacts of road national investments, it has some practical issues. First, as our investment measure is based on road lengths crossing municipal boundaries, large area municipalities might overestimate highway investments. For instance, a municipality can receive a wide road segment, but its economic center be far from the nearest road. Whether it occurs, we could observe and highway investment overestimation. This problem is critical in the North and Mid-West regions. Second, as exposed by several studies (Calderón and Servén, 2014; Straub, 2011), monetary measures as ours are more likely to be constrained by institutional issues as inefficiency, corruption, and harmful bureaucracy. In this sense, a suitable identification strategy is needed at measuring the causal highway investment impact on local outcomes using this data.

As robustness checks, we will also try two additional road variables. The first one is a dummy variable assuming value one if the municipality received a road investment during the PAC period, and zero otherwise. The second one is the road length growth rate between 2006 and 2018. In this case, we use 2006 data from the 2007 National Transport Logistics Plan (PNLT) and 2018 data from DNIT[28].

[27] We cannot fully compare our data with the one from Medeiros *et al.* (2021b) as they accounted for maintenance expenses. However, they compared the road investments annual average of R$ 12.31 billion in the PAC period against an annual average of R$ 5.26 billion between 1995 and 2006.

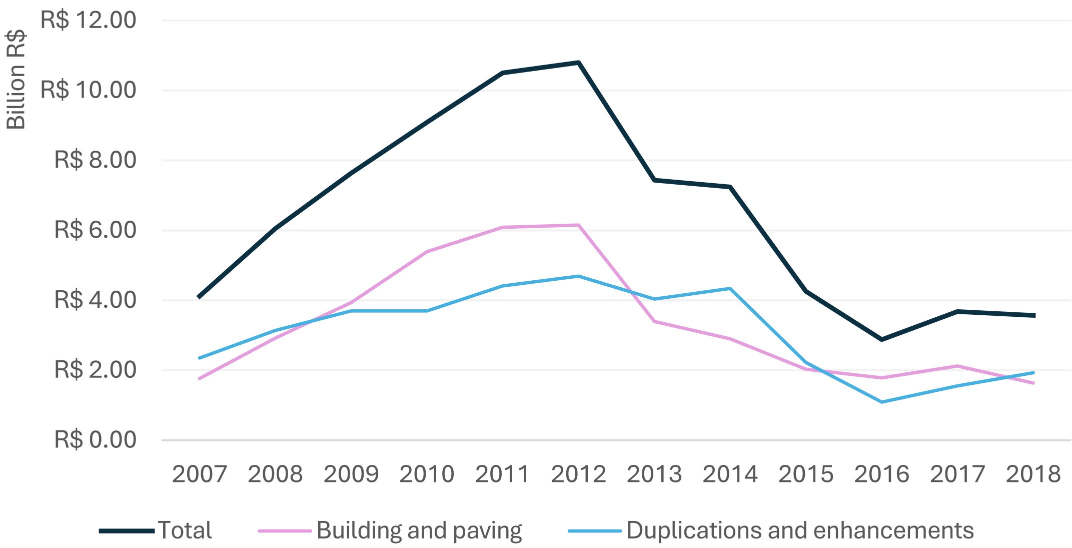
[28] We can likely observe measurement error in the road length variable as well, as the PNLT and DNIT files are not fully comparable. In addition, there is methodological variations over the years in relation to road classifications as federal, state level and so forth. Then, this variable should be used with caution.

Figure D2. Federal Highway Investments by Municipality



Source: authors' elaboration. Black lines are state boundaries.

Figure D3. PAC Highway Investments (2022 R$ billion): Georeferenced Municipal Data



Source: authors' elaboration. Notes: the values do not account for investments in maintenance.

E. Infrastructure project cost-related IVs

To construct the cost-related IVS we use a set of variables representing environmental, geographical and expropriation costs at the municipal level. The next subsections describe in detail the proposed variables.

1. Environmental costs

At the environmental scope, we use georeferenced data of legal protected areas[29] available in the National Registry of Conservation Units (CNUC), maintained by the Ministry of the Environment (MMA). Then, we merge this data with the municipalities boundaries shapefile to identify whether a legal protected area intersects a municipality. Our variable is a dummy which assumes value 1 if the municipality is intersected by a legal protected area, and 0 otherwise.

Second, we utilize the environmental embargo terms data of the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) inspection system. Environmental embargoes represent penalties applied to prevent an exploratory activity from continuing. The embargos also serve to prevent ongoing damage and promote environmental recovery. For instance, penalties are applied to prevent activities with potential to damage the environment, such as deforestation, pollution, and hunting. The application of environmental embargoes by IBAMA occurs mainly in cases where the degradation or damaging activity involves permanent legal protected areas. To generate our variable, we first aggregate the number of embargoes in the five previous years (2002-2006)[30] from the PAC by municipality. Hence, we create a dummy variable which assumes value 1 if there was an environmental embargo in the municipality during this period, and 0 otherwise.

We also try the share of forest area in relation to the total municipality area in 2006 as a third option. This variable came from MAPBIOMAS (Souza *et al.*, 2020) land cover and land use data, which can be obtained at the municipal level. Whilst this variable can capture an important cost of infrastructure building, it may lack variation as Brazil has an enormous forest area. In this sense, weak instrument issues might be expected.

It is important to point out that the institutions accountable for the data are utterly independent from municipalities. The legal protected areas data are administered by the Brazilian (federal) environmental protection system and are controlled by the Chico Mendes Institute for Biodiversity Conservation (ICMBio), as part of the National System of Nature Conservation Units (SNUC). Similarly, IBAMA is a federal

---

[29] The data are divided into six groups: Federal Full Protection and Sustainable Use Conservation Units, State Full Protection and Sustainable Use Conservation Units, and Municipal Full Protection and Sustainable Use Conservation Units.

[30] We take the 5-years sum to avoid potential outliers at using annual data.

agency related to the MMA. In this sense, concerns about the environmental IVs exogeneity condition are quite reduced.

The rationale behind the environmental IVs is that they impact the feasibility and the efficiency of a highway investment. Conditional on controls, we can expect environmental issues causing delays and rising infrastructure projects costs, which in turn may (indirectly) affect outcome variables.

2. *Geographical costs*

To construct geographic-related IVs we rely on a few studies measuring infrastructure effects on local outcomes using this kind of identification approach (Holl, 2012; Lu *et al.*, 2022; Martin-Barroso, Nunez-Serrano, and Velazquez, 2015; Medeiros *et al.*, 2022; Zhang, Hu, and Lin, 2020). All those works utilize some measure based on slope, elevation, ruggedness, or altitude.

Our preferred measure is the share of the municipality area with slope above 20%, which corresponds to hilly areas. This variable is highly related to road construction in the world and in Brazil, as DNIT defines maximum values for slope to be applied to the construction of highways and local roads. Slope is characterized by the relation between a gradient and a corresponding distance in a very small scale, which is unlikely to affect any development outcome directly at the aggregated municipal level. To calculate this variable, we use slope raster data from the National Institute for Space Research (INPE), which allows us to count the number of slope pixels above the 20% cutoff. Then, we generate the share of hilly pixels in relation to the total pixels as our main geographic IV.

We could try several other geographical variables as additional IVs. The first candidate is the municipal average altitude. The second is an elevation-based measure, calculated as the percentage of non-plain areas in the total municipality area (Lu *et al.*, 2022). Both variables are made available by INPE. Other variables could be average rain, temperature, the proportion of mass water in the total municipal area, longitude, and latitude. All those variables are available in public dataset from the IBGE and the National Information System on Water Resources (SNIRH) of the National Water Agency (ANA).

However, it is important to note that all those potential IVs are much more likely to violate the exclusion restriction than our preferred slope measure. Geographical variables measured as average altitude, elevation, rain or even temperature has been used to instrumentalize several other independent variables. The most common case is institutions, in which we can cite the study of Iasco-Pereira, Romero and Medeiros (2021) using latitude, longitude, temperature, rainfall and altitude as IVs to municipal institutional quality in a growth equation like ours. If geographical IVs affects economic growth through pathways other than highway building, exclusion restriction is likely to be violated (Felton and Stewart, 2022). In addition, whilst the slope variable has a strict relationship with highway building norms in Brazil, some of the other variables may suffer from excess of generality. For instance, even in a high (low) altitude or elevation municipality, there exist unlimited (limited) areas with adequate slope to construct highways.

Like the environmental cost IVs, we expect an indirect impact of geographic costs on outcome variables through the infrastructure measure. Geology complexity tends to raise the level of road investments as well as its efficiency. On the other hand, geographical cost is more easily predicted in the study design phase of the infrastructure project, which might imply hilly regions to be avoided by planners following engineering guidelines. Then, the relationship between geographical instruments will depend on the opposite forces of the higher costs in geographically complex areas and the ability of planners to avoid those areas.

3. Expropriation and interferences (human physical) costs

To quantify expropriation (human) costs we use urban infrastructure building, demographic, and land conflict variables. In the urban context, our preferred variable is the share of urban infrastructure[31] building in relation to the total municipality area. For this, we use land use and land cover data made available by MAPBIOMAS (Souza *et al.*, 2020), extracted at the municipal level. Our second measure is populational density, measured as the number of inhabitants divided by the area (km^2) of the municipality in 2000. The source is the 2000 Demographic Census of the Brazilian Institute of Geography and Statistics (IBGE).

Those urban related demographic variables represent only a part of the problem. It is likely that policymakers will avoid highly dense areas precisely to avoid economic costs of expropriating households and properties. Most of the federal highway extension run through rural areas. However, these areas can also present different cost levels related to human expropriation and interference issues. As seen in Section 3, there are localities where the population have invaded highway segments claiming their land rights. In this sense, there are also additional costs arising from expropriation and interference in the context of rural and isolated regions.

To represent expropriation and interferences costs in the rural context, we use data from the CEDOC Dom Tomás Balduíno of the Pastoral Land Commission (CPT) on land conflicts. More specifically, we get data on households involved in land conflicts in 2006. From that, we aggregate the number of households involved in land conflicts and create a dummy variable which assumes value 1 if the municipality present one

[31] This variable is related to construction and infrastructure and is used to generated urban density areas measures.

or more land conflicts, and 0 otherwise. This variable is a proxy for expropriation and interferences costs in the rural context, as it captures the propensity of a municipality to be impacted by some kind of dispute in rural areas. We argue that the greater the number of households involved in conflicts in the rural context, the greater tends to be the expropriation costs and the larger tends to be the propensity of the population to interfere in potential road buildings crossing the municipality.

All expropriation and interference costs are likely to violate the exclusion restriction for some reasons. First, urban agglomeration – as sometimes measured as population density, which in turn is highly correlated with urban infrastructure building as well – can promote economic growth. Second, rural land conflicts might come from disputes on land rights but also involve economic interests as extractivism and deforestation. If it occurs, we might observe a direct human expropriation and interference effect on economic outcomes. It increases our concern on exclusion restriction and makes critically necessary a careful model specification including several control variables to block potential instrument-outcome confounding.

The rationality of these IVs is that physical human costs – as expropriation and interferences – will demand more resources and time to build road projects, conditional on controls. In this way, they also affect outcome variables trough infrastructure. In this case, however, we need to be further careful in including control variables as physical human costs are expected to be highly correlated with demographic variables. To the best of our knowledge, urban infrastructure has not been tested as an urban cost-related instrument before as well as our land conflict variable is a novel measure to expropriations and interferences in the rural context.

F. Non-Random Road Allocation IVs

We use three sets of instruments to represent the "local *propensity to receive road interventions".* The first ones are constructed using the LCP-MST method. The second ones are based on the Brasília Plan. The third one is based on traffic intensity data.

1. LCP-MST IVs

To construct our LCP-MST hypothetical road system, we first need to point out the hubs that the network is supposed to connect through a minimization cost process (Dijkstra, 1959; Kruskal, 1956)[32]. Our preferred strategy considers the starting and ending points of highways receiving PAC investments as hubs. To identify those central cities, we use SNV georeferenced data and generate a dummy variable assuming value 1 if a municipality is a starting or an ending point of a PAC intervention road, and zero otherwise. For instance, the BR-262 presents the municipalities of Cariacica (ES) and Corumbá (MS) as starting and ending points. Cariacica is part of the Metropolitan Region of Vitória (ES), the capital of the state of Espírito Santo (ES), a coastal area containing one of the most important ports of the country (the Vitória Port). On the other hand, Corumbá is a hinterland municipality in the Mid-West region bordering Bolívia. In addition, Corumbá is one of the most important and richer cities of the Mato Grosso do Sul (MS) state. This example elucidates how the starting and ending points approach might raise hubs of different kinds. Corumbá seems to represent a mixture of a political hub – due to its proximity to an international border – and economic hub – due to its economic importance to the regional economy. In a similar rationality, we can identify several other (74 ending or starting points) hubs related to economic, political, touristic or a combination of those factors and others. This approach gains relevance specially in country contexts wherein road policies have no clear direction, being hard to predict which places governments aimed to connect.

Second, we also try to stablish hubs based on the centrality of cities. For this end, we use the Regions of Influence of Cities (REGIC/IBGE) survey of 2007. The REGIC classifies municipalities in order of its influence on other cities in terms of goods and services provided – for instance, the existence of universities, airports, health facilities and so forth. Then, municipalities are classified into five classes: metropolis; regional capitals; sub-regional centers; zone centers; and local centers. Local centers are municipalities with less influence on others, while metropolis are the most central Brazilian cities. In 2007, the two up classes were represented by 75 (1,4%) municipalities out of 5,280 municipalities evaluated, a similar number

[32] Following Faber (2014), we use the Dijkstra's (1959) optimal route algorithm to compute least costly construction paths between any bilateral pair of targeted nodes. Then, we use these bilateral cost parameters in combination with Kruskal's (1956) minimum spanning tree algorithm to identify the subset of routes that connect all targeted nodes on a single continuous graph subject to global network construction cost minimization.

compared to our ending and starting point hubs. To better illustrate the survey hierarchy, Vitória (ES) is classified as regional capital (second class) and Corumbá (MS) as zone center (fourth class). Based on the REGIC, we create a hub dummy variable assuming value 1 if the municipality is a metropolis or a regional capital, and zero otherwise. It is important to note that this approach considerably modifies the considered hubs in comparison with using the starting and ending points of intervention roads, suggesting that this survey captures more local related influence that might be far from federal government aims. In addition, this approach is more likely to exclude potential political hubs from being hubs if they are more regionally isolated, as the survey is based on urban centrality and the linkages among cities.

Third, we use population historical data from IBGE to stablish which we call "historical" hubs. We create a dummy variable assuming value 1 if a municipality had population above 50,000 in 1920, and zero otherwise. This condition is observed for 98 municipalities, then considered hubs. If current infrastructure policies are just following a path dependence dynamic - represented by historical variables as population in 1920- aiming to serve historically important cities with restrained demand for transportation infrastructure, the considered highly populated historical cities are obvious hub candidates. In this case, we obviously observe a concentration of hubs in coastal areas, as those were first populated due to its access facilities and proximity to ports.

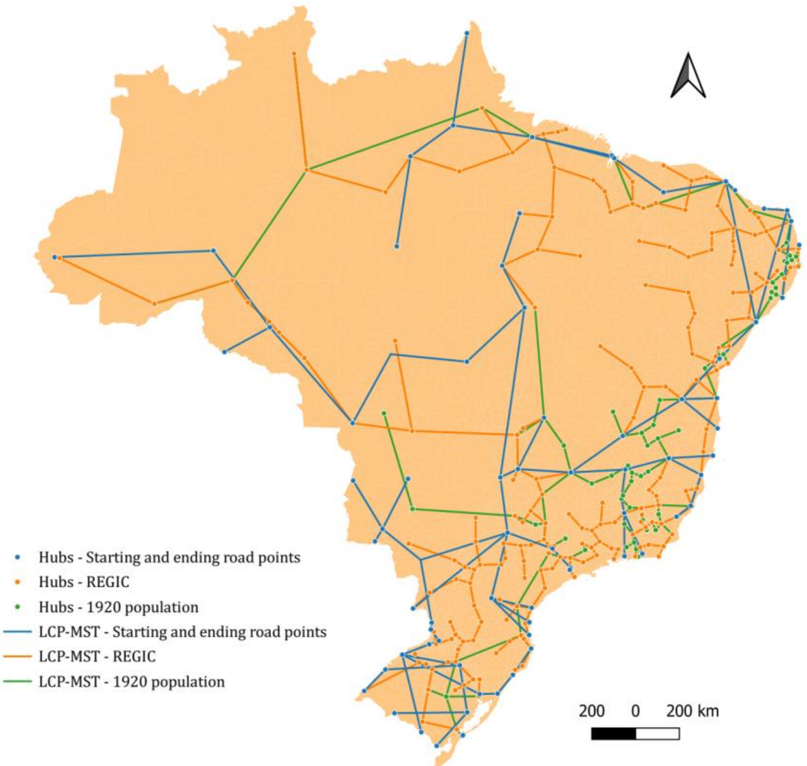
Next, we use the LCP-MST approach to generate hypothetical road networks minimizing the costs of connecting all the stablished hubs for each one of our three proposed groups of hubs. First, we generate simpler hypothetical networks based on Euclidean distance (see Figure F1). Then, we improve the Euclidean distance-based hypothetical networks by including our preferred geographical, environmental, and human physical costs in the LCP-MST minimization cost path.

To create an infrastructure project cost weight, we rely on the ANTT resolutions discussed in earlier sections. ANTT stablished several geographical, environmental, and human physical indicators, among others. For each one of those cost-related measures, the ANTT methodology provided weights capturing their importance on the feasibility of transport projects. We filter the indicators related to geographical, environmental, and human physical costs, and recalculate the importance (weight) of each cost on project feasibility. Environmental costs received a weight of 23.43%, geographical costs 42.68%, and human physical costs 33.89%. Then, we multiply our preferred cost-related measures – sloped area for geographical costs, a weighted averaged of legal protected areas and embargos for environmental costs, and a weighted average of urban infrastructure and land conflicts for human physical costs – of each cost type by their respective ANTT weights. Our cost index ranges from 0 to 1, and the higher its value the more it costs to build roads. Finally, the cost-related weights are interacted with the Euclidean distance and this interaction is used at calculating the LCP-MST network for each one of three potential hubs approaches.

It is important to note the that those novel infrastructure cost-related hypothetical networks are potential instruments for correcting (or alleviating) both measurement error and omitted variable bias from non-random allocation of road investments, constituting an important contribution of this work. We also try additional cost measures as robustness checks.

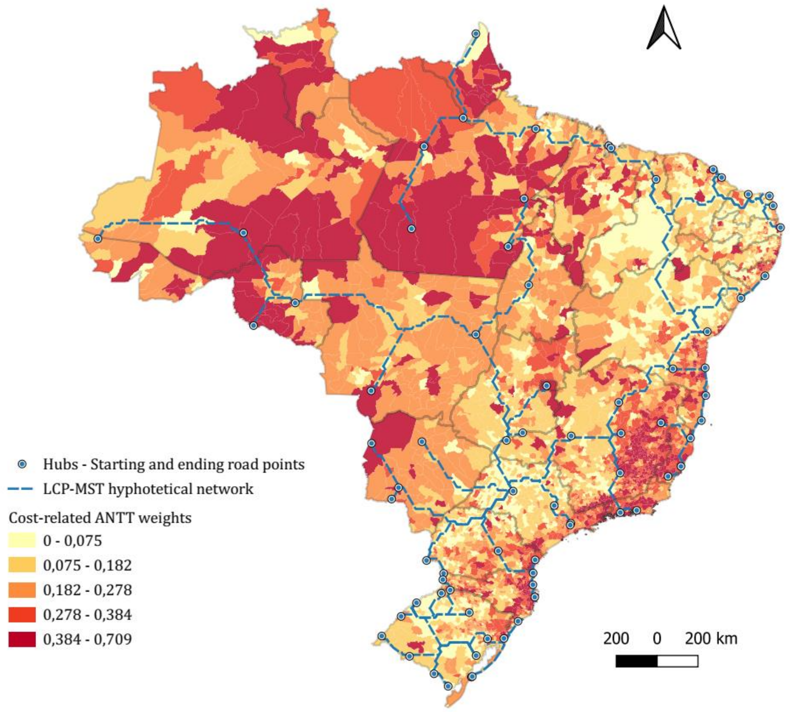
From the hypothetical networks, we create our non-random allocation LCP-MST instruments. The instruments are measured as the log form of the distance from the municipality center to the nearest hypothetical road segment. Figure F2 shows the cost-related ANTT LCP-MST network using starting and ending road points as hubs. REGIC and Historical networks can be seen in Figures F3 and F4.

Figure F1. LCP-MST hypothetical road networks: Euclidean distance



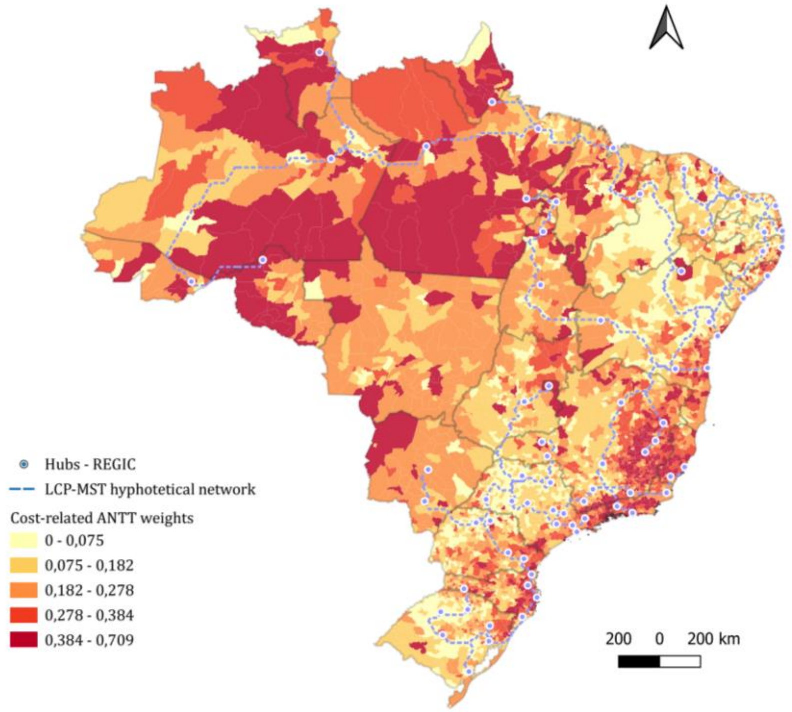
Source: Authors' elaboration.

Figure F2. LCP-MST hypothetical road networks: starting and ending road points using cost-related weights in the cost minimization path



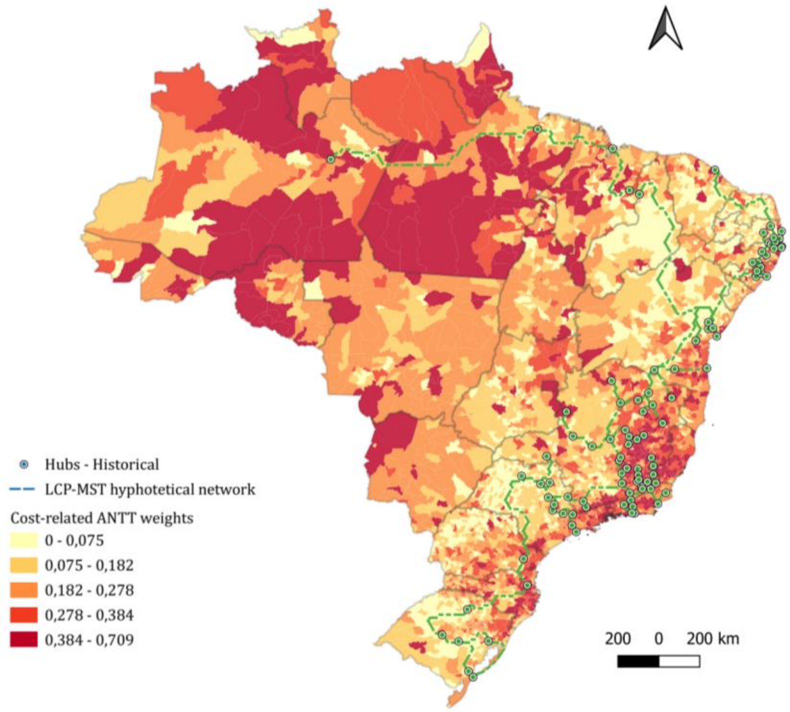
Source: authors' elaboration.

Figure F3. LCP-MST hypothetical road networks: REGIC hubs using cost-related weights in the cost minimization path



Source: Authors' elaboration.

Figure F4. LCP-MST hypothetical road networks: Historical hubs using cost-related weights in the cost minimization path



Source: Authors' elaboration.

2. Political IVs

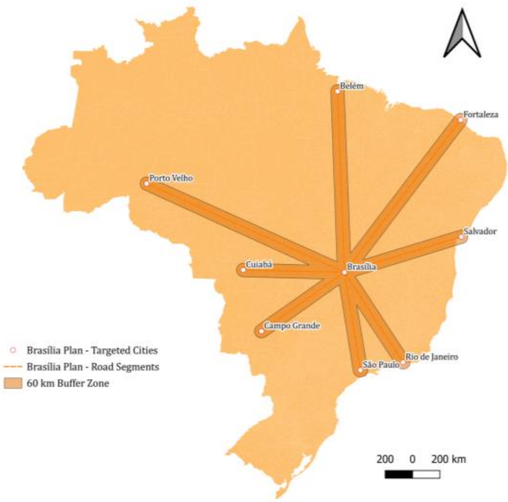
We rely on Bird and Straub (2020) to construct political IVs based on historical plans. Bird and Straub constructed a hypothetical radial (straight line) network connecting the capital Brasília to eight important cities around the country. By linking the capital to those cities, the radial network established corridors, which incidentally connected other localities along the way. The rationality behind this instrument in that, conditional on controls, the Brasília Plan was designed to attend quite different (political) purposes than modern productivity growth.

Our first political IV is measured as the distance from a municipality center to the nearest Brasília Plan segment. We replicate the same index proposed by Bird and Straub (2020) based on buffer zones around the straight lines and the shares of each municipalities' area within each zone.

Our second political IV is based on the Juscelino Kubitschek (JK) Road Cruise, an extension of the Brasília Plan. We digitalized old maps to construct our second historical instrument. The JK Road Cruise's main goals were the same described by Bird and Straub (2020) for the Brasília Plan, i.e., to serve the future Brazilian capital and to connect the whole country to spread the governmental power on the territory as well as to expand the consumption of domestically produced goods and services (Brasil, 1956a, 1956b, 1958). The JK Road Cruise included more hubs than the

Brasília Plan, likely indicating the political bias at developing some specific regional areas not directly connected to Brasília. We follow the procedure in Bird and Straub (2020) and apply it to the JK Cruise network, generating 60 km buffer zones around the lines. Then our two political IVs are distance-based indices from the municipalities centers to the lines weighted by the municipal area shares into the buffer zones. Figure F5 exhibits our two historical networks based on the Brasília Plan.

Figure F5. Historical IVs: the Brasília Plan



(a)

(b)

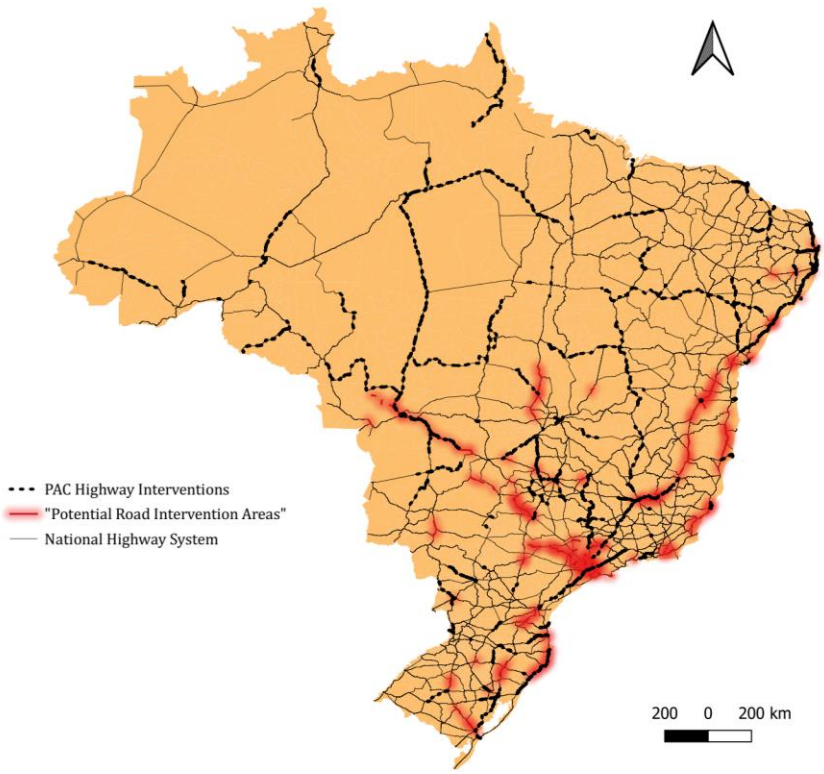
Source: (a) Bird and Straub (2020); (b) authors' elaboration.

3. *"Potential Road Intervention Areas" IV*

An additional econometric issue at using highway investment variables is related to the intervention heterogeneity. The vast empirical literature on infrastructure, growth and productivity has mainly examined the impact of the construction of new roads (Foster *et al.*, 2023a, 2023b). However, in our data, around a half of the total federal highway investment was devoted to duplications and enhancements in localities already connected by the national road network in the PAC starting year. To overcome this issue, we calculate an original instrument based on traffic intensity data in federal roads.

First, we use 2007 PNLT data on traffic intensity to identify "*potential road intervention areas*". In this data, road segments are based on the classification A-F of traffic intensity commonly used in transportation engineering studies in Brazil. Based on the vehicles flows in federal roads, highways segments are classified from A (light traffic) to F (heavy traffic, including huge traffic jams). Then, we consider as heavy traffic roads those segments classified as D, E and F, being of potential federal government focus. Figure F6 shows the "*potential road intervention areas*".

Figure F6. Potential Road Intervention Areas



Source: authors' elaboration.

Next, our instrument is calculated as the distance from the municipality center to the nearest heavy traffic road segment. The rationale behind this IV is that, conditional on controls, municipalities already connected by federal roads in 2006 and located close, but not so close to "*potential road intervention areas*" are more likely to (inconsequentially) receive highway investments to reduce traffic levels in the critical areas. In this case, including demographic controls is a critical condition to ensure IV validity, as we can expect heavy traffic roads crossing urban agglomerations being directly impacted by events involving people, as accidents paralyzing roads, need for traffic signals reducing cars travel speed and so forth.

In Figure F6 we can visualize two patterns relating PAC road interventions and the "*potential road intervention areas*". First and obvious, some heavy traffic areas were targeted by the PAC. In this case, we can rely on the inconsequential unit approach and exclude aimed cities. Second, there are interventions close, but not so close to the "*potential road intervention areas*", likely indicating that highway investments were in some part (incidentally) directed to surrounding cities to reduce traffic congestion in the critical points.

G. Control Variables

Our main dependent variable is the Gross Domestic Product (GDP) *per capita*. We made this choice based on the related literature in which we observe a massive use of GDP as one of the interest variables[33]. In addition, GDP *per capita* can be used as a *proxy* for productivity, which allow us to calculate the return rate to highway investments and provide a more interpretable result for policy purposes. As robustness checks, we also use employment, firms, and wages as labor market measures.

To avoid omitted variable bias, we include an extensive set of controls. Our control variables are based on the same group of studies described previously in Figure 2.1. We also include some controls to capture specificities related to the Brazilian context.

First, we include the initial level of the dependent variable as a control for the municipality development level. Municipalities with different levels of development can present different returns to infrastructure investments. In addition, this variable captures convergence effects (Cosci and Mirra, 2017). It is an important control as policymakers may act to promote balanced economic growth or to foster national growth by targeting developed regions. We also include the share of poor people to control for policies oriented to poverty alleviation, which is a characteristic of the most part of the PAC period.

Second, as road infrastructure is served by the federal and state level governments, we include state fixed effects to control for regional infrastructure policies. This variable also controls for other types of state fixed effects - as institutional, geographical, environmental, and political. This control is also relevant because the PAC realized several investments in road maintenance. However, this road investment data is only available at the state level and cannot be disaggregated by municipality. In this sense, this fixed effect controls for omitted variable bias of other types of PAC road investments.

Third, municipality area is included to control for territorial size. This controls is critical as our main infrastructure variable is based on the PAC road length crossing a municipality area. Then, including this variable is important to not confusing the true highway investment impact on outcomes with the highway impact in a municipality that has just a large area, then receiving more investments because it has a longer highway extension. In addition, this variable also accounts for the possibility that smaller/larger places may have other systematic differences such as institutional quality (Bird and Straub, 2021). As an additional municipal size

[33] About 69% of the studies listed in Figure 2.1 used GDP, GDP *per capita* or some productivity measure as one of the main dependent variables.

variable, we include the formal workforce measured as the log of the number of formal workers in 2007.

Fourth, we include one important control related to the agricultural sector. The PAC period coincided with a period of basic products prices appreciation. If governments aimed at meeting the existing demands of their growing economic sectors, investments in infrastructure may have been partly due to the performance of the agriculture sector in the period (Medeiros *et al.*, 2021b). Then, municipalities specialized in agriculture may have been more likely to receive federal road investments. In this sense, we use the agriculture share in GDP in 2007 as a control.

Fifth, infrastructure investments are seen as an important determinant of exports performance (Coşar and Demir, 2016; Duranton *et al.*, 2014). Exporting municipalities may have received a higher priority in the allocation of national infrastructure investments in the period of good performance in the international market. Then, we include the share of exports of each municipality in the national exports in 2007 as a control.

Sixth, we include a set of controls related to complementary infrastructures. Municipalities well connected with federal and state road networks as well as port and rail infrastructures might be able to benefit more from national road investments. For example, municipalities served with high quality state road network can observe a greater impact of federal highways investments, as these roads may complement other well-established infrastructures. Then, we include the distance (km) to the nearest port, railroad, and state road in 2006 as controls.

Seventh, we control for the historical propensity to a municipality to receive federal road investments. Our variable is the number of railway stations in 1920, the main transportation modal in that period. This variable controls for the propensity of municipalities to receive highway investments because they are historically well located in the country's transportation network. This control is particularly important to the validity of our physical human (interferences and expropriation) IVs as well as our LCP-MST IV based on historical hubs, as they are correlated with demographic and historical variables. In other words, controlling for this variable also alleviates the problem of non-random allocation of federal investments based on past infrastructure. We also control for the distance to the capital Brasília to overcome political related policy aims.

Finally, we include some controls related to the municipal social and institutional background. Those controls are relevant, specially to avoid validity issues on our geographical and expropriation IVs. Geography is the key determinant of climate and natural resource endowments, and it can also play a fundamental role in the disease burden, transport costs, and the extent of diffusion of technology from more advanced areas that societies experience. It therefore exerts a strong influence on agricultural productivity and the quality of human resources (Rodrik, Subramanian and Trebbi, 2004). In addition, if expropriation and interferences are related to land conflicts influenced by economic activity as extractivism, a direct effect of it on economic growth may exist. In this sense, we first include the Index of Municipal Institutional Quality (IQIM)[34] to control for municipal institutions. Second, we include the population share with master or doctoral degree as a proxy for local social development. Table G1 summarizes our variables and their respective data sources. Table G2 shows descriptive statistics.

[34] The IQIM index was constructed considering three dimensions: first, the degree of participation of the population in the public administration, considering capacity of decision, equality, deliberation, and influence in resource allocation; second, the financial capacity of each municipality, the degree to which the funds are generated by the local economic activity without federal resource transfers; and third, the management capacity of local government. The IQIM was normalised to stay between zero and one. The closer to one the index is, the better the institutions are (Iasco-Pereira, Romero, and Medeiros, 2021).

Table G1. Variables description

Type	Variable	Description	Source
Highway	Highway Investments	PAC Highway Investments (R$)	MINFRA
	Highway Intervention	1 if the municipality received PAC highway investments and 0 otherwise	MINFRA
	Highway Length	Highway length (km) growth between 2006 and 2018	PNLT (2007) and DNIT
IVs	Legal Protected Area	1 if the municipality is intersected by a legal protected area and 0 otherwise	MMA
	Environmental Embargos	1 if there was an environmental embargo in the municipality during this period and 0 otherwise	IBAMA
	Forest area	Forest area (km^2)/Total area (km^2)	
	Slope	Area with slope above 20% (which corresponds to hilly areas) (km^2)/Total area (km^2)	INPE
	Altitude	Average altitude (m)	INPE
	Non-plain areas	Non-plain areas/Total area (km^2)	INPE
	Urban infrastructure	Building and infrastructure area (km^2)/ Total area (km^2)	MAPBIOMAS (Souza *et al.*, 2020)
	Population density	Population/Area (km^2)	IBGE
	Land conflicts	1 if there exist a household involved in land conflicts and 0 otherwise	CEDOC - CPT
	Cost Index 1	First component of the MCA	-
	Cost Index 2	Second component of the MCA	-
	LCP-MST Starting and Ending Road Points	Distance to the nearest LCP-MST hypothetical line using starting and ending road points as hubs	MINFRA
	LCP-MST REGIC	Distance to the nearest LCP-MST hypothetical line using REGIC first three classes as hubs	REGIC
	LCP-MST Historical	Distance to the nearest LCP-MST hypothetical line using historical cities (cities with population above 50,000 in 1920) as hubs	IBGE
	Brasília Plan	Distance to the nearest Brasília Plan line (weighted by the municipality area share into the buffer zone)	Bird and Straub (2020)
	JK Road Cruise	Distance to the nearest JK Road Cruise line (weighted by the municipality area share into the buffer zone)	Brasil (1956)
	Potential Road Intervention Area	Distance to the nearest road segment classified as heavy traffic (D, E or F classification)	PNLT (2007)
	Non-Random Allocation Index	First component of the PCA	-
Dependents	GDP *per capita*	Gross Domestic Product (R$)/Population	RAIS and IBGE
	Wages	Wages (R$)	RAIS/MTE
	Firms	Number of firms	RAIS/MTE
	Employment	Number of workers	RAIS/MTE
Controls	GDP *per capita*, lagged	Gross Domestic Product (R$)/ Number of workers in 2007	IBGE
	Share of poor people (%)	Population below the poverty line/Total population	
	Area	Municipality area (km^2)	IBGE
	Work force	Number of formal workers	RAIS/MTE
	Agriculture share (%)	Agriculture Value Added (R$)/ Total Value Added (R$)	IBGE
	Exports share (%)	Municipal Exports (US$) / National Exports (US$)	MDIC
	Distance to state road	Distance (km) to the nearest state road	MINFRA
	Distance to railroad	Distance (km) to the nearest railroad	MINFRA
	Distance to port	Distance (km) to the nearest federal port	MINFRA
	Railways stations in 1920	Number of railways stations in 1920	Rede Ferroviária Federal S/A
	Distance to Brasília	Distance (km) to the capital Brasília	
	Institutional Quality	Institutional quality municipal index (IQIM)	Ministry of Planning
	Human Capital (%)	Workers with master of doctoral degree/Total workers	RAIS/MTE

Source: Author's elaboration. Note: IBGE - Brazilian Institute of Geography and Statistics; ANA - National Water and Sanitation Agency; BCB - Central Bank of Brazil; INPE - National Institute for Space Research; IBAMA - Brazilian Institute of Environment and Renewable Natural Resources; MDIC - Ministry of Development, Industry, Commerce and Services; MINFRA - Ministry of Infrastructure; PNLT - National Transport Logistics Plan; MT – Ministry of Transport; RAIS - Annual Social Information Report; MTE - Ministry of Labor and Employment; SIM - Mortality Information System; MS – Ministry of Health.

Table G2. Descriptive statistics

Type	Variable	Obs	Mean	Std. Dev.	Min	Max
Highway	Highway Investments	5,570	1.273	3.449	0.000	14.772
	Highway Intervention	5,570	0.125	0.331	0.000	1.000
IVs	Legal Protected Area	5,570	0.424	0.494	0.000	1.000
	Environmental Embargos	5,570	0.385	0.487	0.000	1.000
	Forest area	5,568	0.377	0.261	0.000	0.994
	Slope	5,565	0.134	0.177	0.000	0.745
	Altitude	5,507	412.476	293.136	0.000	1628.000
	Non-plain areas	5,575	0.998	0.009	0.749	1.000
	Urban infrastructure	5,555	0.020	0.069	0.000	1.000
	Population density	5,565	3.134	1.417	-2.029	9.454
	Land conflicts	5,575	0.063	0.243	0.000	1.000
	Cost Index 1	5,550	0.000	1.141	-1.375	4.578
	Cost Index 2	5,550	0.000	1.020	-2.573	10.767
	LCP-MST Starting and Ending Road Points	5,565	3.706	2.110	-13.816	6.941
	LCP-MST REGIC	5,565	3.825	1.074	-0.112	6.763
	LCP-MST Historical	5,565	4.233	1.307	-1.141	7.344
	Brasília Plan	5,566	4.965	1.635	1.609	7.310
	JK Road Cruise	5,568	4.181	1.648	1.204	7.310
	Potential Road Intervention Area	5,565	4.605	1.452	-3.546	7.670
	Non-Random Allocation Index	5,565	0.000	1.634	-2.279	2.851
Dependents	GDP *per capita*	5,564	0.329	0.346	-2.130	3.083
	Wages	5,563	0.068	0.142	-0.632	7.093
	Firms	5,548	0.178	0.290	-0.301	4.170
	Employment	5,564	0.793	0.498	-2.272	6.529
Controls	GDP *per capita*, lagged	5,564	2.791	0.724	1.096	6.533
	Share of poor people (%)	5,565	41.057	22.776	0.700	90.760
	Area	5,570	6.205	1.279	1.271	11.980
	Work force	5,536	6.049	2.169	0.000	15.023
	Agriculture share (%)	5,564	0.220	0.153	0.000	0.839
	Exports share (%)	5,570	0.000	0.001	0.000	0.034
	Distance to state road	5,565	5.599	37.296	0.001	740.818
	Distance to railroad	5,565	333.165	258.873	0.353	1271.520
	Distance to port	5,565	91.916	215.006	0.031	2081.518
	Railways stations in 1920	5,570	0.536	2.877	0.000	107.000
	Distance to Brasília	5,565	1075.822	445.390	0.000	2872.215
	Institutional Quality	5,505	3.023	0.551	1.000	4.904
	Human Capital (%)	5,564	0.001	0.004	0.000	0.174

H. Correcting measurement error bias and selecting suitable cost-related IVs: the second step

Before we start the econometric analysis, we present some correlations and preliminary tests on our instruments. Figure H1 shows the correlation matrix including the highway investment variables and the proposed instruments. Highway variables are positively associated with environmental and expropriation IVs and negatively with geographical IVs. Obviously, in most cases, IVs correlate more strongly with other IVs of the same type. For instance, urban infrastructure and population density are highly correlated. The same occurs (in lower magnitude) between legal protected areas and environmental embargos or sloped areas and altitude. The lower correlation among cost-related IVs of different types seems to suggest that those variables capture different spectrums of infrastructure project costs. It is important to note that the forest area variable presented a very poor correlation with the highway variables, which places some caution on the following analyses.

Table H1 shows naïve OLS estimations on the relationship between highway investments and cost-related IVs. We include the full set of controls. In Table H2, we include naïve OLS estimations without controls. We also try our highway intervention dummy variable as an additional highway investment variable (Tables H3-H4). The results remain unchanged. Results corroborates the relationships observed in the correlation matrix. As we can see, the estimated coefficients for forest area and populational density are not significant, likely indicating a weak IV issue. In Column 10, we include all IVs together, and results remain unchanged. In Columns 11-15 we include combinations of our preferred instruments for environmental, expropriation and geographical costs. Results suggest that all our preferred cost-related IVs are significantly correlated with the highway investment variables.

In Table H5, we propose a set of specifications based on our preferred cost-related IVs. As we believe different kinds of costs affecting infrastructure investment in different forms and magnitudes, we start by including at least one measure of each cost type by specification. We have not included forest area and population density in Table 1 regressions as they did not influence highway investments. To check instrument strength, we report KP Wald F and *effective* F (Olea and Pflueger, 2013) statistics. Both statistics are robust to heteroskedasticity and weak instruments, and the *effective* F statistic works suitably even in multiple instruments setting (Andrews, Stock and Sun, 2019), as proposed by our identification strategy.

The first stage regressions show that our cost-related IVs are strong predictors of the long-term changes in the national highway investments at the municipal level. Nonetheless, some IVs seem to violate exclusion restriction or unconfoundedness.

Figure H1. Correlation matrix: National Highway investments and potential cost-related IVs

Variables	Infrastructure		Cost-Related IVs											Non-Random Allocation IVs					
			Environmental			Geographical			Human Physical			Aggregated		LCP-MST			Political		Intervention Areas
	Log Highways Investments	Highway Intervention	Legal Protected Areas	Environmental Embargo	Forest Area	Sloped Area	Altitude	Non-Plain Areas	Urban Infrastructure	Populational Density	Land Conflicts	Cost Index 1	Cost Index 2	LCP-MST Start/End Points	LCP-MST REGIC	LCP-MST Historical	Brasília Plan	JK Road Cruise	Intervention Areas
Log Highways Investments	1.00																		
Highway Intervention	0.97	1.00																	
Legal Protected Areas	0.12	0.12	1.00																
Environmental Embargo	0.19	0.17	0.24	1.00															
Forest Area	0.01	0.01	0.14	0.09	1.00														
Sloped Area	-0.08	-0.07	0.11	0.05	0.08	1.00													
Altitude	-0.10	-0.08	-0.05	-0.05	-0.24	0.19	1.00												
Non-Plain Areas	-0.10	-0.08	0.01	-0.03	0.06	0.04	0.17	1.00											
Urban Infrastructure	0.15	0.14	0.08	0.06	-0.12	-0.04	-0.02	-0.16	1.00										
Populational Density	0.06	0.06	-0.06	-0.09	-0.31	0.14	0.04	-0.09	0.58	1.00									
Land Conflicts	0.10	0.09	0.14	0.12	0.09	-0.09	-0.13	0.02	0.00	-0.10	1.00								
Cost Index 1	0.18	0.17	0.76	0.71	0.13	0.36	-0.01	-0.03	0.29	0.08	0.12	1.00							
Cost Index 2	0.17	0.15	-0.03	0.08	-0.14	-0.70	-0.14	-0.15	0.74	0.31	0.07	0.00	1.00						
LCP-MST Start/End Points	-0.12	-0.10	-0.02	-0.03	0.17	0.02	0.09	0.08	-0.20	-0.26	0.00	-0.06	-0.16	1.00					
LCP-MST REGIC	-0.11	-0.11	-0.05	-0.08	0.07	-0.03	0.02	0.00	-0.16	-0.31	-0.10	-0.12	-0.10	0.12	1.00				
LCP-MST Historical	-0.08	-0.07	-0.07	-0.10	0.09	-0.11	-0.03	-0.03	-0.19	-0.29	-0.05	-0.16	-0.07	0.16	0.28	1.00			
Brasília Plan	-0.02	-0.03	-0.06	-0.11	-0.04	0.10	-0.22	-0.09	-0.12	0.03	0.01	-0.09	-0.16	-0.11	-0.03	0.02	1.00		
JK Road Cruise	-0.05	-0.05	-0.01	-0.05	-0.09	-0.02	-0.07	0.13	-0.21	-0.19	0.04	-0.08	-0.14	0.00	0.03	0.11	0.40	1.00	
Intervention Areas	-0.08	-0.08	0.01	0.00	0.42	-0.26	-0.26	0.10	-0.25	-0.44	0.14	-0.12	-0.01	0.21	0.09	0.17	0.01	0.16	1.00

Source: author's elaboration.

Table H1. Highway Investments and Cost-Related IVs: Naive OLS Regressions

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Legal Protected Areas	0.390***									0.350***	0.434***		0.423***		
	(0.10)									(0.10)	(0.10)		(0.10)		
Environmental Embargos		0.372***								0.276**		0.374***		0.367***	
		(0.11)								(0.11)		(0.11)		(0.11)	
Forest Area			-0.037							0.325					
			(0.24)							(0.26)					
Sloped Area				-1.270***						-1.126***	-1.364***	-1.199***	-1.362***	-1.201***	-1.141***
				(0.33)						(0.35)	(0.33)	(0.33)	(0.33)	(0.33)	(0.33)
Altitude					-0.001***					-0.001***					
					(0.00)					(0.00)					
Non-plain areas						-32.862***				-22.882**					
						(9.40)				(9.28)					
Urban Infrastructure							3.981***			3.316**	3.455***	3.534***	3.454***	3.530***	3.686***
							(1.28)			(1.37)	(1.29)	(1.29)	(1.29)	(1.29)	(1.30)
Populational Density								-0.041		-0.135					
								(0.14)		(0.14)					
Land conflicts									0.561**	0.486**			0.525**	0.542**	0.557**
									(0.25)	(0.25)			(0.25)	(0.25)	(0.25)
Constant	-5.289***	-5.361***	-5.597***	-5.262***	-5.668***	26.971***	-4.451***	-5.572***	-5.553***	18.795**	-3.887***	-4.021***	-3.853***	-3.981***	-4.185***
	(1.32)	(1.31)	(1.32)	(1.31)	(1.31)	(9.44)	(1.33)	(1.31)	(1.31)	(9.28)	(1.34)	(1.33)	(1.33)	(1.32)	(1.33)
Observations	5468	5468	5466	5468	5468	5468	5456	5468	5468	5455	5456	5456	5456	5456	5456
R^2 Adjusted	0.148	0.148	0.146	0.148	0.154	0.153	0.148	0.146	0.147	0.165	0.153	0.152	0.154	0.153	0.151

All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table H2. Highway Investments and Cost-Related IVs: Naive OLS Regressions without controls

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Legal Protected Areas	0.853***									0.527***	0.853***		0.765***		
	(0.10)									(0.09)	(0.10)		(0.10)		
Environmental Embargos		1.320***								1.100***		1.298***		1.228***	
		(0.10)								(0.10)		(0.10)		(0.10)	
Forest Area			0.087							-0.074					
			(0.17)							(0.16)					
Sloped Area				-1.579***						-1.473***	-1.734***	-1.651***	-1.558***	-1.503***	-1.289***
				(0.21)						(0.22)	(0.21)	(0.21)	(0.21)	(0.21)	(0.20)
Altitude					-0.001***					-0.001***					
					(0.00)					(0.00)					
Non-plain areas						-33.504***				-22.878**					
						(8.83)				(9.55)					
Urban Infrastructure							7.325***			5.424***	6.679***	6.627***	6.759***	6.684***	7.214***
							(1.17)			(1.39)	(1.15)	(1.13)	(1.15)	(1.13)	(1.17)
Populational Density								0.125***		0.064					
								(0.05)		(0.05)					
Land conflicts									1.505***	0.910***			1.159***	1.078***	1.396***
									(0.26)	(0.25)			(0.26)	(0.25)	(0.26)
Constant	0.911***	0.765***	1.241***	1.483***	1.750***	34.721***	1.130***	0.881***	1.178***	23.592**	1.013***	0.865***	0.952***	0.803***	1.216***
	(0.05)	(0.05)	(0.08)	(0.06)	(0.09)	(8.82)	(0.05)	(0.15)	(0.05)	(9.52)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Observations	5570	5570	5568	5565	5507	5570	5555	5565	5570	5493	5550	5550	5550	5550	5550
R^2 Adjusted	0.015	0.035	-0.000	0.006	0.010	0.008	0.021	0.002	0.011	0.078	0.041	0.060	0.048	0.066	0.036

Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table H3. Highway Investments and Cost-Related IVs: Naive Logit Regressions without controls

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Legal Protected Areas	0.731***									0.518***	0.750***		0.691***		
	(0.08)									(0.09)	(0.08)		(0.08)		
Environmental Embargos		1.005***								0.864***		1.010***		0.964***	
		(0.08)								(0.09)		(0.08)		(0.09)	
Forest Area			0.155							0.105					
			(0.15)							(0.16)					
Sloped Area				-1.335***						-1.453***	-1.501***	-1.424***	-1.366***	-1.303***	-1.111***
				(0.25)						(0.27)	(0.25)	(0.26)	(0.26)	(0.26)	(0.26)
Altitude					-0.001***					-0.000**					
					(0.00)					(0.00)					
Non-plain areas						-14.920***				-11.203***					
						(3.82)				(4.14)					
Urban Infrastructure							3.616***			1.923***	3.177***	3.236***	3.252***	3.296***	3.566***
							(0.52)			(0.63)	(0.49)	(0.49)	(0.50)	(0.50)	(0.54)
Populational Density								0.113***		0.101**					
								(0.04)		(0.04)					
Land conflicts									0.913***	0.507***			0.660***	0.631***	0.854***
									(0.13)	(0.14)			(0.14)	(0.14)	(0.14)
Constant	-2.301***	-2.419***	-2.000***	-1.783***	-1.618***	12.946***	-2.034***	-2.307***	-2.019***	8.422**	-2.212***	-2.335***	-2.257***	-2.382***	-1.972***
	(0.06)	(0.06)	(0.07)	(0.05)	(0.07)	(3.81)	(0.04)	(0.13)	(0.04)	(4.12)	(0.07)	(0.07)	(0.07)	(0.07)	(0.05)
Observations	5570	5570	5568	5565	5507	5570	5555	5565	5570	5493	5550	5550	5550	5550	5550
R^2 Adjusted															

Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table H4. Highway Investments and Cost-Related IVs: Naive Logit Regressions

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Legal Protected Areas	0.339***									0.326***	0.399***		0.394***		
	(0.10)									(0.10)	(0.10)		(0.10)		
Environmental Embargos		0.201*								0.122		0.205*		0.201*	
		(0.11)								(0.11)		(0.11)		(0.11)	
Forest Area			0.221							0.559**					
			(0.26)							(0.28)					
Sloped Area				-1.328***						-1.478***	-1.505***	-1.278***	-1.506***	-1.282***	-1.252***
				(0.39)						(0.40)	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)
Altitude					-0.001***					-0.001***					
					(0.00)					(0.00)					
Non-plain areas						-17.157***				-11.308***					
						(4.98)				(4.28)					
Urban Infrastructure							1.767***			1.339*	1.342**	1.434**	1.341**	1.432**	1.514**
							(0.62)			(0.70)	(0.63)	(0.62)	(0.63)	(0.63)	(0.63)
Populational Density								-0.089		-0.133					
								(0.11)		(0.11)					
Land conflicts									0.242	0.195			0.219	0.236	0.244
									(0.17)	(0.17)			(0.17)	(0.17)	(0.17)
Constant	-6.683***	-6.727***	-6.866***	-6.484***	-6.856***	10.030**	-6.276***	-6.766***	-6.847***	5.570	-5.744***	-5.852***	-5.728***	-5.831***	-5.959***
	(1.09)	(1.09)	(1.10)	(1.11)	(1.09)	(5.02)	(1.12)	(1.10)	(1.10)	(4.33)	(1.12)	(1.12)	(1.12)	(1.12)	(1.13)
Observations	5467	5467	5465	5467	5467	5467	5455	5467	5467	5454	5455	5455	5455	5455	5455
R^2 Adjusted															

All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table H5. Federal Highway Investments and Municipal GDP *per capita* Growth, 2007-2018: Second Step 2SLS IV Regressions

	1	2	3	4	5	6	7	8	9	10	11
Second stage											
Log Highways Investments	0.0287***	0.0243**	0.0074	0.0024	0.0251***	0.0336***	0.0415***	0.0347**	0.0259***	0.0339***	0.0240***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
First stage											
Legal Protected Areas	0.4580***		0.2859***	0.3914***	0.4478***	0.4997***	0.5096***		0.4255***	0.4719***	
	(0.10)		(0.10)	(0.10)	(0.10)	(0.10)	(0.10)		(0.10)	(0.10)	
Environmental Embargos		0.3876***						0.4292***	0.3390***	0.3721***	
		(0.11)						(0.11)	(0.11)	(0.11)	
Sloped Area	-1.4738***	-1.2936***			-1.4695***	-1.7850***	-1.7880***	-1.5968***	-1.5083***	-1.8155***	
	(0.33)	(0.33)			(0.33)	(0.33)	(0.33)	(0.33)	(0.33)	(0.33)	
Altitude			-0.4280***								
			(0.06)								
Non-plain Areas				-28.3482***							
				(9.77)							
Urban Infrastructure	5.0701***	5.2007***	5.0793***	5.0443***	5.0867***				4.9258***		
	(1.22)	(1.22)	(1.19)	(1.23)	(1.22)				(1.21)		
Land Conflict					0.5528**	0.5469**					
					(0.25)	(0.25)					
Cost Index 1											0.0160***
											(0.00)
Cost Index 2											0.0329***
											(0.01)
Observations	5466	5466	5331	5466	5466	5478	5478	5478	5466	5478	5466
KP Wald F Statistic	21.128	17.582	28.017	15.729	17.215	18.282	24.594	18.772	18.012	19.699	27.275
Effective F Statistic	20.857	18.677	29.628	14.354	16.429	15.538	24.599	18.671	18.795	20.054	33.915
2SLS critical value for tau=5%	21.250	21.256	22.085	25.517	23.039	17.848	3.238	4.208	22.554	14.091	14.888
2SLS critical value for tau=10%	13.299	13.292	13.639	15.597	13.984	11.258	3.118	3.636	13.806	9.009	9.888
R^2	0.17	0.19	0.24	0.23	0.19	0.15	0.10	0.14	0.19	0.15	0.19

All regressions include the following set of control variables: GDP per capita in 2007; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

In columns 1 and 2 we include our preferred measures of each infrastructure cost type. In both cases, *effective* F statistics are not large enough when compared to the 2SLS critical value for τ equal to 5%, but they are using τ equal to 10%. Next, we try several other IV combinations in columns 3-10.

In columns 3 and 4 we replace sloped area for altitude and non-plain areas, respectively. In column 5 we include the land conflict variable together with urban infrastructure, whilst in column 6 we exclude urban infrastructure and maintain land conflict. The low *effective* F statistic in columns 4, 5 and 6 points to weak instruments issues and likely estimation bias. We then investigate the reasons for those issues. In addition, the non-significant and quite smaller highway parameters in columns 3 and 4 raises concerns in using altitude and non-plain areas as IVs for highway investments.

Regarding geographical IVs, our main suggestion is the inadequacy of the altitude – at least as it is measured, being the municipal average altitude – and the non-plain area variables as IVs. Geography variables, as average altitude, are used as controls in income equations as ours to capture the direct geography impact (through agricultural productivity, for instance) and the indirect impact (through the institutional quality channel) on economic growth. In our estimations, the direct impact seems to be non-negligible – and, whether it occurs, geography affects both income and highway investments, and estimation bias is expected. This finding cautions against using the average altitude measure as IV in studies about infrastructure and development in underdeveloped countries as Brazil. On the other hand, our preferred measure (sloped area) is the proportion of hilly areas in relation to the total municipal area, which means that even municipalities located in high (low) average altitude areas might have a large (small) proportion of their area geographically suitable for road investments at lower costs. The differences between these two geographical variables can be seen in the far from high correlation (0.18) between them (Figure A1), suggesting that they are predicting geographical issues quite distinctly. In addition, the non-plain area presents quite small variation (mean around 0.99 and standard deviation around 0.01), as Brazil is characterized by small parcels of non-plain areas, and this lack of variation might be affecting econometric results.

On the expropriation and interferences variables, results put caution about using this kind of infrastructure project costs as IV. Urban infrastructure seems to be a suitable but slightly weak IV for highway investments, as its inclusion in the model lowers *effective* F statistics below the critical values in some specifications (it can be seen comparing columns 10 and 9, for instance). Even conditional on several controls, urban infrastructure is likely impacting economic growth directly. Similarly, land conflicts seem to have a non-negligible direct effect on growth, which is likely related to economic related land disputes as extractivism.

Next, in columns 7 and 8 we try combinations of our preferred environmental and geographical cost IVs, as they seem to be more plausibly exogenous in comparison to the human physical costs. Both specifications are satisfactory, presenting *effective* F statistics suitable at the 5% level. In column 9 we include urban infrastructure together with our suitable IVs based on columns 7 and 8. We can note that, when including the urban infrastructure measure, *effective* F statistic decreases to a level smaller than the 2SLS critical values when choosing τ equal to 5%. This finding corroborates the results obtained in the previous columns and raises some caution at using expropriation and interferences measures as IVs to road investments.

Finally, we also try some composite cost indexes (Cost Index 1 and Cost Index 2) based on dimensionality reduction methods[35], as cost types may have some complementary characteristics. In addition, working with a smaller number of instruments might alleviate weak IV and overidentification concerns. We generate our composite indexes using our preferred measures – legal protected areas, environmental embargos, sloped area, and urban infrastructure. Cost Index 1 can be interpreted as an environmental cost index, whilst Cost Index 2 can be seen as a geographical-expropriation index. Results in column 11 shows a significant effect of both cost indexes on highways and *effective* F statistics considerably increases.

Then, we turn our analysis to the IVs parameters signals and magnitude. Environmental IVs positively affect federal highway investments. This result was expected as more environmental costly places might demand a higher level of investments *per* kilometer of road. Similarly, the higher the expropriation and interferences costs, the higher the value of investments in federal highways tends to be demanded. On the other hand, geographical costs negatively affect highway investments. One possible explanation is that geographic costs can be more easily predicted from engineering studies in the design phase of the infrastructure project. In this sense, its negative sign may represent the fact that these regions are being avoided by planners. On the other hand, expropriation costs have a highly unpredictable aspect as they are impacted by legal disputes involving people. In the same way, environmental costs might raise additional bureaucracy due to environmental licenses not initially planned in the infrastructure project design, as well as lead to new legal disputes involving native populations claiming for their rights on legally protected areas.

Regarding the signal and magnitude of the highway investment parameter in the second stage, it corroborates an extensive number of previous studies which found a positive highway investment impact on local outcomes. Based on our specifications with suitable *effective*-F statistics, we find a second step highway investment elasticity between 0.02 and 0.03, which is in the range between 0 and 0.06 found by Foster *et al.* (2023b) based on a huge number of infrastructure econometrics studies.

[35] We use Multiple Correspondence Analysis (MCA) for mixed data to generate our composite indexes. We use the first two components as they accumulate 59% of the original data variation. We name the first component as Cost Index 1, and it is calculated by $Cost\ Index\ 1 = 0.58 * Legal\ Protected\ Area + 0.51 * Embargos + 0.13 * Sloped\ Area + 0.08 * Urban\ Infrastructure$. The second component, which we name Cost Index 2, is calculated by $Cost\ Index\ 2 = 0.00 * Legal\ Protected\ Area + 0.01 * Embargos + 0.49 * Sloped\ Area + 0.58 * Urban\ Infrastructure$. We expect the Cost Index 1 to be more plausibly exogenous, as it received lower influence from the urban infrastructure variable.

In Foster *et al.* paper, the average elasticity for the transportation sector is 0.03, which brings our estimates even closer to the basis.

Finally, we can discuss the direction of the measurement error bias by comparing our estimations in Table 1 with OLS regressions (see Table H6). As we expected, the OLS elasticities (around 0.004) are downward biased. Whereas inefficiencies and delays due to geographical, environmental, and human physical issues occur, measurement error takes place by inflating the highway investment variable. In other words, more economic resources are needed to construct or improve a kilometer of road that will be finally used by the population. This measurement error leads to an underestimation of the strength of the (positive) association between transportation infrastructure and economic activity.

By applying our second step identification strategy, we fixed (or alleviated) the expected measurement error underestimation bias. In this sense, our second step parameters may be understood as a kind of "*free from measurement error*" elasticity. Nonetheless, two obvious empirical issues remain. First, if our instruments are not fully or truly capturing the main infrastructure costs leading to inefficiencies, some measurement error bias might remain. This issue seems more troublesome for human physical IVs. Second, non-random allocation bias might exist even after correcting for measurement error in highway variable. The next section presents several tests on the latter bottleneck.

Table H6. Federal Highway Investments and Municipal GDP *per capita* Growth, 2007-2018: OLS Regressions

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Log Highways Investments	0.0006	0.0004	0.0078***	0.0069***	0.0041***	0.0043***	0.0043***	0.0042***	0.0040***	0.0040***	0.0040***	0.0036***	0.0035**	0.0035**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log GDP 2007			-0.2792***	-0.2823***	-0.3227***	-0.3370***	-0.3390***	-0.3435***	-0.3444***	-0.3446***	-0.3446***	-0.3661***	-0.3721***	-0.3721***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Log Municipal area				0.0258***	0.0136***	0.0043	0.0042	0.0004	-0.0029	-0.0031	-0.0035	0.0126**	0.0134***	0.0134***
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Log employment					0.0269***	0.0396***	0.0394***	0.0407***	0.0424***	0.0425***	0.0427***	0.0275***	0.0268***	0.0268***
					(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Agriculture share (% GDP)						0.2852***	0.2890***	0.2780***	0.2876***	0.2872***	0.2881***	0.2898***	0.2774***	0.2774***
						(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Exports by municipality (% total)							5.3664	5.4599	5.6395	5.6710	5.7111	7.2610	7.4979	7.4931
							(4.64)	(4.61)	(4.62)	(4.63)	(4.63)	(4.70)	(4.84)	(4.84)
Distance to Brasilia								-0.0001***	-0.0002***	-0.0002***	-0.0002***	-0.0001**	-0.0001***	-0.0001***
								(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance to the nearest state road									0.0007***	0.0007***	0.0007***	0.0006***	0.0006***	0.0006***
									(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance to the nearest port										0.0000	0.0000	0.0000	0.0000	0.0000
										(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance to the nearest railroad											0.0000	0.0000	0.0000	0.0000
											(0.00)	(0.00)	(0.00)	(0.00)
Share of poor people (%)												-0.0042***	-0.0044***	-0.0044***
												(0.00)	(0.00)	(0.00)
Number of railway stations in 1920													-0.0013	-0.0013
													(0.00)	(0.00)
Institutional Quality													0.0198**	0.0198**
													(0.01)	(0.01)
Graduate education (% workers)														0.0396
														(0.72)
Constant	0.3280***	0.4168***	1.1915***	0.9970***	1.0422***	1.0023***	1.0088***	1.2527***	1.3167***	1.3100***	1.2939***	1.3994***	1.3633***	1.3633***
	(0.00)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.08)	(0.08)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)
Observations	5564	5564	5564	5564	5535	5535	5535	5535	5535	5535	5535	5535	5478	5478
State FE	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R^2 Adjusted	-0.000	0.071	0.237	0.242	0.256	0.265	0.265	0.268	0.271	0.271	0.271	0.284	0.285	0.285

Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

I. Solving non-random allocation bias: the third step

Table I1. Federal Highway Investments and Municipal GDP *per capita* Growth, 2007-2018: 2SLS IV Regressions

	1	2	3	4	5	6
Second stage						
Log Highways Investments	0.0109**	0.0109**	0.0147***	0.0144***	0.0097*	0.0096**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
First stage						
Legal Protected Areas	0.4795***		0.4877***		0.4387***	
	(0.10)		(0.10)		(0.10)	
Environmental Embargos	0.3503***		0.3530***		0.3369***	
	(0.11)		(0.11)		(0.11)	
Sloped Area	-1.4095***		-1.5196***		-1.6325***	
	(0.31)		(0.31)		(0.31)	
Cost Index 1		-0.4253***		0.2654***		0.2326***
		(0.02)		(0.05)		(0.05)
Cost Index 2		0.2683***		0.3676***		0.3668***
		(0.05)		(0.07)		(0.07)
LCP-MST REGIC	-0.4213***	0.3602***				
	(0.02)	(0.07)				
LCP-MST Historical			-0.3840***	-0.3869***		
			(0.02)	(0.02)		
JK Road Cruise					-0.3312***	-0.3312***
					(0.02)	(0.02)
Observations	5402	5391	5402	5391	5402	5391
KP Wald F Statistic	118.735	162.994	116.089	158.400	112.011	151.930
Effective F Statistic	111.454	122.519	111.443	121.939	96.225	103.699
LIML critical value for tau=5%	19.954	21.214	19.696	20.974	20.885	22.720
R^2	0.23	0.23	0.22	0.22	0.23	0.23

All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

J. Robustness checks

In this section, we present several robustness checks to our main results showed in the previous two topics. First, we propose a falsification test based on a novel planned road sample exercise. Second, we try other commonly used dependent variables. Third, we test two different measures of road infrastructure. Fourth, we run additional robustness check using the limited information maximum likelihood (LIML) estimator, bootstrap standard errors, and excluding potential outliers.

1. Falsification test

If our instruments are valid, they should affect the outcome only through the highway investment variable. Therefore, cost-related and non-random road allocation IVs should have no effect on local economic outcomes that are not in the highway investment pathway. A way to test the exclusion restriction is using falsification tests based on alternative study samples (Felton and Stewart, 2022; Pizer, 2016). A useful falsification sample would not be exposed to the treatment (highway investment), but it would be subjected to all the potential confounders that might be correlated with the instrument and the outcome, like demographic, economic, complementary infrastructure, and institutional features.

To test the exclusion restriction, we specify a sample closely related to our study population but not receiving highway investments. First, we select all municipalities which had federal planned roads in 2006. We extracted this data from the same 2007 PNLT shapefile. Second, we use the 2018 SNV shapefile to compare[36] whether those planned roads in 2006 were still classified as planned in 2018. Next, we excluded those municipalities that presented planned roads for both years and received PAC investments, as it could capture an inefficiency or a long-time to build infrastructure issue.

Following the DNIT road terminology, planned highways are those roads that do not physically exist yet. Practically, planned roads represent a hypothetical highway acting as a guideline intended to meet a potential traffic demand. In this sense, our rationality behind this falsification test is that places (hypothetically) crossed by planned roads were similarly demanding road infrastructure interventions, likely exposed to the same potential confounders as the full sample. As none of those municipalities received federal highway interventions, the falsification test is performed by including the IVs in an alternative specification of the outcome equation. Table J1 summarizes the results. We run the same specifications of Table 2 using the planned roads sample and include our preferred cost-related IVs separately as well (columns 1 and 2).

---

[36] It is important to note that those two datasets are not fully comparable, as methodological changes occurred and some measurement error is expected. Then, results need to be taken with caution.

Table J1. Falsification test: OLS regressions

	1	2	3	4	5	6	7	8	9	10	11	12
Legal Protected Areas	0.0256		0.0267		0.0272		0.0262		0.0250		0.0269	
	(0.03)		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
Environmental Embargos	0.0349		0.0344		0.0349		0.0334		0.0351		0.0342	
	(0.03)		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
Sloped Area	0.0243		0.0217		0.0198		0.0207		0.0215		0.0204	
	(0.10)		(0.10)		(0.10)		(0.10)		(0.10)		(0.10)	
Cost Index 1		0.0219		0.0222		0.0226		0.0217		0.0219		0.0222
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)
Cost Index 2		0.0066		0.0073		0.0078		0.0072		0.0073		0.0075
		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)
LCP-MST Starting and Ending Road Points			-0.0040	-0.0040					0.0164	0.0135		
			(0.01)	(0.01)					(0.02)	(0.02)		
Potential Road Intervention Area					0.0061	0.0056			0.0112	0.0098		
					(0.01)	(0.01)			(0.02)	(0.02)		
Brasília Plan							-0.0054	-0.0049	-0.0074	-0.0062		
							(0.01)	(0.01)	(0.02)	(0.02)		
Non-random Allocation Index											-0.0081	-0.0075
											(0.01)	(0.01)
Observations	403	400	403	400	403	400	403	400	403	400	403	400
R^2	0.41	0.40	0.41	0.40	0.41	0.41	0.41	0.40	0.41	0.41	0.41	0.40

All regressions include the following set of control variables: GDP per capita in 2007; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

The results show that GDP *per capita* growth is not correlated with our proposed cost-related IVs in the falsification sample between 2007 and 2018[37]. Conditional on controls, results suggest that our suitable cost-related and non-random allocation IVs do not violate exclusion restriction. These results increase confidence on our identification strategy.

2. *Outcomes*

In this section, we estimate the federal highway investments impact on employment, firms, and wages. This test is important to validate our empirical strategy and to confirm the highway investment influence on local outcomes. We use the econometric specification in column 10 of Table 2 and apply it to each one of the dependent variables using formal labor market data from RAIS/MTE.

Table J2 shows the results. For all tested dependent variables, the cost-related and non-random allocation IVs are strong predictors of national highway investments. In addition, the signal and significance of the first stage coefficients remained quite similar, as well as remains positive the road impact on outcomes. Our results suggest that our identification strategy is suitable not only for the GDP *per capita* variable, but that it might also be suitable for predicting the highway investment impacts on several other labor market local outcomes.

Table J2. Federal Highway Investments and Municipal Outcomes Growth, 2007-2018: 2SLS IV Regressions

	1	2	3
Second stage	ΔEmployment	ΔFirms	ΔWages
Log Highways Investments	0.0168***	0.0214***	0.0697***
	(0.00)	(0.00)	(0.01)
First stage			
Cost Index 1	0.2207***	0.2376***	0.2133***
	(0.05)	(0.05)	(0.05)
Cost Index 2	0.3808***	0.4250***	0.3683***
	(0.07)	(0.07)	(0.07)
Non-random allocation Index	-0.5069***	-0.5113***	-0.5056***
	(0.02)	(0.03)	(0.02)
Observations	5141	5127	5141
KP Wald F Statistic	146.472	149.102	145.535
Effective F Statistic	114.103	117.330	112.904
2SLS critical value for tau=5%	20.686	20.826	20.609
R^2	0.12	0.47	0.20

All regressions include the following set of control variables: GDP per capita in 2007; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

[37] In unreported estimations, the exceptions are urban infrastructure and populational density variables, which presented significant coefficients in all specifications even controlling for past GDP, population, distance to complementary infrastructures and past infrastructure. This result seems to indicate that expropriation and interferences IVs in the urban context are likely violating exclusion restriction. On the other hand, the land conflict IV, which is related to the rural context, appears to be uncorrelated with GDP *per capita* growth in the falsification sample.

3. Highway investment measure

Several studies correctly argues that monetary variables - such as the investment flows used in our study - may contain several measurement errors (Calderón and Servén, 2014; Kenny, 2009; Straub, 2011). To increase our empirical strategy reliability, we also try a dummy (intervention) measure of road infrastructure. Related studies have used similar measures to capture access to highways (Asher and Novosad, 2017; Frye, 2016; Michaels, 2008; Percoco, 2015; Zhang, Hu and Lin, 2020). Our dummy variable assumes value 1 if the municipality was crossed by a PAC highway intervention, and zero otherwise. If there are too high measurement error in our preferred investment flow variable, the intervention dummy variable might alleviate the problem as it does not contain monetary values anymore. On the other hand, the intervention variable gives the same weight to all municipalities receiving highway investments, which is a not negligible empirical issue. Second, we try a road length variable following a vast strand of literature (Baum-Snow *et al.*, 2020; Duranton *et al.*, 2014; Foster *et al.*, 2023a, 2023b; Straub, 2011). Tables J3 and J4 summarizes the results using the same specifications following Table 2.1. As expected, the highway intervention parameters are positive and significant in all specifications, and the same holds for road length. The cost-related and non-random allocation IVs work in the same way as at using continuous highway investment flows, corroborating previous estimates.

4. Additional robustness checks

To raise confidence on our main estimates, we run the same specifications of Table 2.1 using the LIML estimator (Anderson and Rubin, 1949). In the overidentification scenario as proposed by our identification strategy, 2SLS estimates might be biased towards OLS as the bias is proportional to the degree of overidentification (Angrist and Krueger, 2001). In overidentified models, LIML is approximately unbiased in the sense that the median of its sampling distribution is generally close to the population parameter being estimated. In addition, we also provide bootstrap confidence intervals. Young (2022) finds that bootstrapped confidence intervals perform better in real-world settings as heteroscedasticity and weak IV assumptions are likely violated. Tables J5 and J6 exhibits the results. In both cases findings remain unchanged, indicating that our main estimations are reliable.

Next, we try several specifications excluding potential outliers. In this set of estimations, we use column 10 of Table 2.1 as our benchmark specification. Results are described in Table J7. In column 1, we drop all municipalities of the state of São Paulo. São Paulo is the richest Brazilian state, representing more than 30% of the national GDP. In addition, São Paulo presents the best road infrastructure of the country, which is in substantial part privately managed. The major share of those high-quality roads is responsibility of the São Paulo state government, being federal roads a small fraction of the total. The result might be a small fraction of the PAC highway investment directed to a high road demanding state, and an underestimation bias could be expected. In column 2, we exclude all municipalities pertaining to North region states. Those municipalities are characterized by large territorial areas, which might bias our highway measure variable as it depends on the road length crossing the municipalities. In column 3, we exclude both São Paulo and Northern municipalities. In column 4, we consider highway investment values smaller than R$ 50 million to be zero. This test is important as we relied on road length crossing municipal areas to construct our highway flow measure, and short road segments (and consequentially small investments values) might be poorly capturing a highway intervention. Finally, in column 5 we exclude municipalities in the top 1 and bottom 1 percentiles of GDP *per capita* growth. In general, findings remain almost unchanged. The most noticeable variation comes from the exclusion of São Paulo municipalities. In columns 1 and 3, the elasticity is around 0.017, whilst our benchmark estimate is 0.012. This result suggest that municipalities of São Paulo might be slightly downward biasing our estimates.

Table J3. Federal Highway Intervention and Municipal GDP *per capita* Growth, 2007-2018: 2SLS IV Regressions

	1	2	3	4	5	6	7	8	9	10
Second stage										
Highway Intervention	0.1613***	0.1594***	0.1236**	0.1206**	0.1057**	0.1065**	0.1231**	0.1215**	0.1262**	0.1239**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
First stage										
Legal Protected Areas	0.0445***		0.0332***		0.0405***		0.0322***		0.0325***	
	(0.01)		(0.01)		(0.01)		(0.01)		(0.01)	
Environmental Embargos	0.0208**		0.0235**		0.0228**		0.0226**		0.0226**	
	(0.01)		(0.01)		(0.01)		(0.01)		(0.01)	
Sloped Area	-0.0940***		-0.1309***		-0.0997***		-0.1175***		-0.1185***	
	(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
Cost Index 1		0.0227***		0.0185***		0.0215***		0.0178***		0.0179***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Cost Index 2		0.0264***		0.0349***		0.0280***		0.0301***		0.0303***
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
LCP-MST Starting and Ending Road Points	-0.0392***	-0.0396***					-0.0129***	-0.0129***		
	(0.00)	(0.00)					(0.00)	(0.00)		
Potential Road Intervention Area			0.0326***	0.0330***			0.0091*	0.0102**		
			(0.00)	(0.00)			(0.00)	(0.00)		
Brasília Plan					-0.0316***	-0.0317***	-0.0135***	-0.0128***		
					(0.00)	(0.00)	(0.00)	(0.00)		
Non-random Allocation Index									-0.0503***	-0.0507***
									(0.00)	(0.00)
Observations	5402	5391	5190	5178	5469	5457	5126	5115	5126	5115
KP Wald F Statistic	114.906	159.088	96.673	137.599	124.326	172.444	72.896	90.639	107.042	148.520
Effective F Statistic	104.916	128.927	96.316	118.452	107.287	131.457	74.150	85.131	105.074	125.647
2SLS critical value for tau=5%	19.156	18.280	18.468	17.721	19.684	19.487	24.055	24.703	19.174	17.705
R^2	0.22	0.22	0.24	0.24	0.23	0.23	0.24	0.24	0.23	0.24

All regressions include the following set of control variables: GDP per capita in 2007; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table J4. Federal Highway Length and Municipal GDP *per capita* Growth, 2007-2018: 2SLS IV Regressions

	1	2	3	4	5	6	7	8	9	10
Second stage										
Highway Length	0.0599***	0.0618***	0.0392*	0.0402*	0.0413*	0.0423**	0.0464**	0.0481**	0.0433**	0.0449**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
First stage										
Legal Protected Areas	0.0631**		0.0412		0.0553**		0.0407		0.0376	
	(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
Environmental Embargos	0.0735**		0.0816***		0.0842***		0.0747**		0.0742**	
	(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
Sloped Area	-0.0595		-0.0824		-0.0950		-0.0639		-0.0433	
	(0.08)		(0.08)		(0.08)		(0.08)		(0.08)	
Cost Index 1		0.0473***		0.0397***		0.0441***		0.0386***		0.0374***
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
Cost Index 2		0.0284**		0.0287**		0.0278*		0.0265*		0.0216
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
LCP-MST Starting and Ending Road Points	-0.0962***	-0.0963***					-0.0365***	-0.0371***		
	(0.01)	(0.01)					(0.01)	(0.01)		
Potential Road Intervention Area			0.0834***	0.0836***			0.0682***	0.0690***		
			(0.01)	(0.01)			(0.02)	(0.02)		
Brasília Plan					-0.0701***	-0.0700***	0.0159	0.0171*		
					(0.00)	(0.00)	(0.01)	(0.01)		
Non-random Allocation Index									-0.1205***	-0.1204***
									(0.01)	(0.01)
Observations	5402	5391	5190	5178	5469	5457	5126	5115	5126	5115
KP Wald F Statistic	78.667	105.397	67.451	90.156	79.975	106.662	49.599	59.166	73.442	97.481
Effective F Statistic	74.122	97.551	73.153	96.950	64.749	85.279	49.766	60.408	71.571	98.108
2SLS critical value for tau=5%	19.242	16.584	19.462	15.656	20.755	18.607	24.240	23.681	19.691	16.136
R^2	0.21	0.21	0.23	0.23	0.23	0.22	0.23	0.23	0.23	0.23

All regressions include the following set of control variables: GDP per capita in 2007; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table J5. Federal Highway Investments and Municipal GDP *per capita* Growth, 2007-2018: LIML IV Regressions

	1	2	3	4	5	6	7	8	9	10
Second stage										
Log Highways Investments	0.0167***	0.0163***	0.0128**	0.0123***	0.0112**	0.0111**	0.0127**	0.0126***	0.0130***	0.0128***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
First stage										
Legal Protected Areas	0.4597***		0.3555***		0.4232***		0.3389***		0.3423***	
	(0.10)		(0.10)		(0.10)		(0.10)		(0.10)	
Environmental Embargos	0.3312***		0.3478***		0.3547***		0.3383***		0.3384***	
	(0.11)		(0.11)		(0.11)		(0.11)		(0.11)	
Sloped Area	-1.3606***		-1.6060***		-1.4378***		-1.4601***		-1.4676***	
	(0.31)		(0.31)		(0.32)		(0.30)		(0.29)	
Cost Index 1		0.2586***		0.2274***		0.2470***		0.2177***		0.2190***
		(0.05)		(0.05)		(0.05)		(0.05)		(0.05)
Cost Index 2		0.3571***		0.4427***		0.3744***		0.3882***		0.3896***
		(0.07)		(0.07)		(0.07)		(0.07)		(0.07)
LCP-MST Starting and Ending Road Points	-0.3875***	-0.3918***					-0.1193***	-0.1204***		
	(0.02)	(0.02)					(0.04)	(0.04)		
Potential Road Intervention Area			0.3252***	0.3294***			0.0893*	0.1025**		
			(0.02)	(0.02)			(0.05)	(0.05)		
Brasília Plan					-0.3170***	-0.3176***	-0.1421***	-0.1314***		
					(0.01)	(0.01)	(0.03)	(0.03)		
Non-random Allocation Index									-0.4994***	-0.5035***
									(0.02)	(0.02)
Observations	5402	5391	5190	5178	5469	5457	5126	5115	5126	5115
KP Wald F Statistic	113.896	156.650	97.221	137.336	123.195	169.934	72.567	89.717	106.925	147.383
Effective F Statistic	104.053	117.056	96.535	108.326	108.318	122.313	73.324	78.795	104.841	112.493
LIML critical value for tau=5%	11.784	19.258	11.535	19.848	11.675	19.872	9.119	14.261	12.006	19.770
R^2	0.22	0.22	0.23	0.24	0.23	0.23	0.23	0.23	0.23	0.23

All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table J6. Federal Highway Investments and Municipal GDP *per capita* Growth, 2007-2018: 2SLS IV Regressions, Bootstrap

	1	2	3	4	5	6	7	8	9	10
Second stage										
Log Highways Investments	0.0167***	0.0163***	0.0128**	0.0123***	0.0112**	0.0111**	0.0127*	0.0126***	0.0130***	0.0128***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
First stage										
Legal Protected Areas	0.4717***		0.3522***		0.5168***		0.2715***		0.3919***	
	(0.10)		(0.09)		(0.09)		(0.09)		(0.10)	
Environmental Embargos	0.3521***		0.2537**		0.4822***		0.2391**		0.4339***	
	(0.10)		(0.11)		(0.11)		(0.11)		(0.11)	
Sloped Area	-1.3929***		-1.5401***		-1.4923***		-1.6424***		-1.1440***	
		0.1414***		0.1988***		0.1725***		0.2889***		0.3117***
Cost Index 1		(0.05)		(0.05)		(0.05)		(0.05)		(0.05)
		0.3435***		0.5802***		0.4647***		0.4671***		0.3779***
Cost Index 2		(0.07)		(0.08)		(0.06)		(0.07)		(0.07)
LCP-MST Starting and Ending Road Points	-0.3643***	-0.3687***					-0.1954***	-0.1582***		
	(0.02)	(0.02)					(0.04)	(0.04)		
Potential Road Intervention Area			0.3293***	0.3150***			0.0421	0.0788		
			(0.02)	(0.02)			(0.05)	(0.05)		
Brasília Plan					-0.3058***	-0.3461***	-0.1296***	-0.1202***		
					(0.01)	(0.02)	(0.04)	(0.03)		
Non-random Allocation Index									-0.5297***	-0.5502***
									(0.03)	(0.02)
Observations	5402	5391	5190	5178	5469	5457	5126	5115	5126	5115
KP Wald F Statistic	113.896	156.650	97.221	137.336	123.195	169.934	72.567	89.717	106.925	147.383
R^2	0.22	0.22	0.23	0.24	0.23	0.23	0.23	0.23	0.23	0.23

All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Bootstrapped standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table J7. Federal Highway Investments and GDP *per capita* Growth, 2007-2018: 2SLS IV Regressions, Additional Robustness Checks

	1	2	3	4	5
Second stage					
Log Highways Investments	0.0165***	0.0124**	0.0172***	0.0103***	0.0118***
	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
First stage					
Cost Index 1	0.2105***	0.1756***	0.1637***	0.4942***	0.2240***
	(0.05)	(0.05)	(0.05)	(0.10)	(0.05)
Cost Index 2	0.4829***	0.3690***	0.4697***	0.7491***	0.3935***
	(0.08)	(0.07)	(0.08)	(0.15)	(0.07)
Non-random allocation Index	-0.5200***	-0.4506***	-0.4649***	-0.5613***	-0.4986***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.02)
Observations	4553	4685	4123	5115	5013
KP Wald F Statistic	139.029	113.102	104.261	63.067	144.673
Effective F Statistic	115.795	81.081	83.184	49.303	108.496
2SLS critical value for tau=5%	19.585	21.902	20.523	24.925	21.311
R^2	0.26	0.23	0.25	0.22	0.20

All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

II. APPENDIX - INFRASTRUCTURE, GROWTH AND REGIONAL DISPARITIES

A. Variables description and descriptive statistics

Table A1. Variables description

Type	Variable	Description	Source
Highway	Highway Investments	PAC Highway Investments (R$)	MINFRA
	Highway Intervention	1 if the municipality received PAC highway investments and 0 otherwise	MINFRA
Road Related Variables	Efficiency	Gross Domestic Product (GDP)(R$)/Road Stock (R$)	IBGE, Frischtak and Mourão (2017) and PNLT (2007)
	Redistribution	Gross Domestic Product (GDP)(R$)/Population	IBGE
	Equity	Road Stock/Area (km^2)	PNLT (2007) and IBGE
	Road Specialization	Share of the intermediate consumption related to the land transportation sector (%)	RAIS/MTE and IBGE (2018)
IVs	Legal Protected Area	1 if the municipality is intersected by a legal protected area and 0 otherwise	MMA
	Environmental Embargos	1 if there was an environmental embargo in the municipality during this period and 0 otherwise	IBAMA
	Slope	Area with slope above 20% (which corresponds to hilly areas) (km^2)/Total area (km^2)	INPE
	Urban infrastructure	Building and infrastructure area (km^2)/ Total area (km^2)	MAPBIOMAS (Souza *et al.*, 2020)
	Population density	Population/Area (km^2)	IBGE
	Cost Index 1	First component of the MCA. Formula: $Cost\ Index\ 1 = 0.58 * Legal\ Protected\ Area + 0.51 * Embargos + 0.13 * Sloped\ Area + 0.08 * Urban\ Infrastructure.$	Medeiros *et al. (2024)*
	Cost Index 2	Second component of the MCA. Formula: $Cost\ Index\ 2 = 0.00 * Legal\ Protected\ Area + 0.01 * Embargos + 0.49 * Sloped\ Area + 0.58 * Urban\ Infrastructure$	Medeiros *et al. (2024)*
	LCP-MST Starting and Ending Road Points	Distance to the nearest LCP-MST hypothetical line using starting and ending road points as hubs	MINFRA
	Brasília Plan	Distance to the nearest Brasília Plan line (weighted by the municipality area share into the buffer zone)	Bird and Straub (2020)
	Potential Road Intervention Area	Distance to the nearest road segment classified as heavy traffic (D, E or F classification)	PNLT (2007)
	Non-Random Allocation Index	First component of the PCA. Formula: $Nonrandom\ Allocation\ Index = 0.58 * \text{LCP} - \text{MST Starting and Ending Road Points} + 0.58 * BrasíliaPlan - 0.57 * PotentialInterventionAreas$	Medeiros *et al. (2024)*
Dependents	GDP *per capita*	Gross Domestic Product (R$)/Population	RAIS and IBGE
	Firms	Number of firms	RAIS/MTE
	Employment	Number of workers	RAIS/MTE
Controls	GDP *per capita*, lagged	Gross Domestic Product (R$)/ Number of workers in 2007	IBGE
	Share of poor people (%)	Population below the poverty line/Total population	
	Area	Municipality area (km^2)	IBGE
	Work force	Number of formal workers	RAIS/MTE
	Agriculture share (%)	Agriculture Value Added (R$)/ Total Value Added (R$)	IBGE
	Exports share (%)	Municipal Exports (US$) / National Exports (US$)	MDIC
	Distance to state road	Distance (km) to the nearest state road	MINFRA
	Distance to railroad	Distance (km) to the nearest railroad	MINFRA
	Distance to port	Distance (km) to the nearest federal port	MINFRA

Railways stations in 1920	Number of railways stations in 1920	Rede Ferroviária Federal S/A
Distance to Brasília	Distance (km) to the capital Brasília	
Institutional Quality	Institutional quality municipal index (IQIM)	Ministry of Planning
Human Capital (%)	Workers with master of doctoral degree/Total workers	RAIS/MTE

Source: Author's elaboration. Note: IBGE - Brazilian Institute of Geography and Statistics; ANA - National Water and Sanitation Agency; BCB - Central Bank of Brazil; INPE - National Institute for Space Research; IBAMA - Brazilian Institute of Environment and Renewable Natural Resources; MDIC - Ministry of Development, Industry, Commerce and Services; MINFRA - Ministry of Infrastructure; PNLT - National Transport Logistics Plan; MT – Ministry of Transport; RAIS - Annual Social Information Report; MTE - Ministry of Labor and Employment; SIM - Mortality Information System; MS – Ministry of Health.

Table A2. Descriptive statistics

Type	Variable	Obs	Mean	Std. Dev.	Min	Max
Highway	Highway Investments	5,570	1.273	3.449	0.000	14.772
	Highway Intervention	5,570	0.125	0.331	0.000	1.000
Road Related Variables	Efficiency	5,570	26.541	405.054	0.000	29,0003.920
	Redistribution	5,570	29.309	27.908	4.373	461.033
	Equity	5,570	7.706	7.333	0.000	102.086
	Road Specialization	5,570	0.041	0.020	0.000	0.107
IVs	Legal Protected Area	5,570	0.424	0.494	0.000	1.000
	Environmental Embargos	5,570	0.385	0.487	0.000	1.000
	Slope	5,565	0.134	0.177	0.000	0.745
	Urban infrastructure	5,555	0.020	0.069	0.000	1.000
	Population density	5,565	3.134	1.417	-2.029	9.454
	Cost Index 1	5,550	0.000	1.141	-1.375	4.578
	Cost Index 2	5,550	0.000	1.020	-2.573	10.767
	LCP-MST Starting and Ending Road Points	5,565	3.706	2.110	-13.816	6.941
	Brasília Plan	5,565	4.965	1.635	1.609	7.310
	Potential Road Intervention Area	5,565	4.605	1.452	-3.546	7.670
	Non-Random Allocation Index	5,565	0.000	1.634	-2.279	2.851
Dependents	GDP *per capita*	5,564	0.329	0.346	-2.130	3.083
	Wages	5,563	0.068	0.142	-0.632	7.093
	Firms	5,548	0.178	0.290	-0.301	4.170
	Employment	5,564	0.793	0.498	-2.272	6.529
Controls	GDP *per capita*, lagged	5,564	2.791	0.724	1.096	6.533
	Share of poor people (%)	5,565	41.057	22.776	0.700	90.760
	Area	5,570	6.205	1.279	1.271	11.980
	Work force	5,536	6.049	2.169	0.000	15.023
	Agriculture share (%)	5,564	0.220	0.153	0.000	0.839
	Exports share (%)	5,570	0.000	0.001	0.000	0.034
	Distance to state road	5,565	5.599	37.296	0.001	740.818
	Distance to railroad	5,565	333.165	258.873	0.353	1271.520
	Distance to port	5,565	91.916	215.006	0.031	2081.518
	Railways stations in 1920	5,570	0.536	2.877	0.000	107.000
	Distance to Brasília	5,565	1075.822	445.390	0.000	2872.215
	Institutional Quality	5,505	3.023	0.551	1.000	4.904
	Human Capital (%)	5,564	0.001	0.004	0.000	0.174

B. Additional estimates

Table B1. Determinants of road infrastructure investments, 2007-2018: OLS Regressions

	1	2	3	4	5	6	7	8	9	10
Efficiency	0.2251***				0.2883***					
	(0.06)				(0.08)					
Redistribution		0.3603***			0.1665					
		(0.12)			(0.13)					
Equity			-0.0356		0.1956**					
			(0.07)		(0.09)					
Road Specialization				0.3179*	0.4022**					
				(0.18)	(0.18)					
Efficiency, Decile Average						0.2672***				0.2726***
						(0.06)				(0.08)
Redistribution, Decile Average							0.2943**			0.1062
							(0.12)			(0.13)
Equity, Decile Average								-0.1235**		0.0316
								(0.05)		(0.07)
Road Specialization, Decile Average									0.3955**	0.4115**
									(0.16)	(0.16)
Observations	5466	5466	5466	5466	5466	5466	5466	5466	5466	5466
R^2	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21

All regressions include the following set of control variables: Cost Index 1, Cost Index 2, Non-Random Allocation Index, GDP per capita in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table B2. Federal Highway Investments and Local Outcomes Growth, 2007-2018: 2SLS IV Regressions

	1	2	3	4
Sample	Efficiency	Redistribution	Equity	Road Specialization
<=50%	0.0294***	0.0124***	0.0092	0.0011
	(0.01)	(0.00)	(0.01)	(0.01)
>50%	0.0123*	0.0054	0.0102	0.0200***
	(0.01)	(0.01)	(0.01)	(0.01)
>50% & <=90%	0.0149**	0.0143***	0.0051	0.0190***
	(0.01)	(0.00)	(0.01)	(0.01)
>50% & <=80%	0.0275***	0.0077*	0.0000	0.0191***
	(0.01)	(0.00)	(0.01)	(0.01)
>10% & <=50%	0.0288***	0.0114***	0.0223**	0.0098
	(0.01)	(0.00)	(0.01)	(0.01)
>20% & <=50%	0.0314***	0.0084	0.0199	0.0121
	(0.01)	(0.00)	(0.01)	(0.01)

All regressions include the following set of control variables: GDP per capita in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table B3. Federal Highway Investments and Local Outcomes Growth by "Ideal" Samples, 2007-2018: 2SLS IV Regressions

	1	2	3	4	5	6	7
Sample	Road Specialization > 50%	Road Specialization>50% & Redistribution <= 50%	Road Specialization > 50% & Equity <= 50%	Road Specialization > 50% & 10% <Equity <= 50%	Road Specialization > 50% & Redistribution <= 50% & Equity <= 50%	Road Specialization > 50% & Redistribution <= 50% & 10% < Equity <= 50%	Road Specialization <= 50%
Efficiency > 50%	0.0246***	0.0143	0.01958*	0.0217	0.0184*	0.0261	-0.0067
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
50%< Efficiency < 90%	0.0273***	0.0148	0.0258**	0.0296**	0.0219*	0.0269	-0.0059
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
50%< Efficiency < 80%	0.0345***	0.0127	0.0302**	0.0438**	0.0212*	0.0141	0.0068
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Redistribution <= 50%	0.0220***	-	0.0165*	0.0316**	-	-	0.0048
	(0.01)		(0.01)	(0.01)			(0.01)
Equity <= 50%	0.0133	-	-	-	-	-	0.0070
	(0.01)						(0.01)
10% < Equity <= 50%	0.0205*	-	-	-	-	-	0.0185
	(0.01)						(0.02)

All regressions include the following set of control variables: GDP per capita in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table B4. Federal Highway Investments and Local Outcomes Growth, 2007-2018: 2SLS IV Regressions, interaction models

	1	2	3
Log Efficiency	0.1132***		
	(0.01)		
Log Redistribution		0.3492***	
		(0.03)	
Log Equity			0.0032
			(0.01)
Log Highways Investments * φ	0.9727***	1.5382***	0.0884
	(0.19)	(0.58)	(0.21)
Log Highways Investments * Log Efficiency * φ	-0.2146***		
	(0.08)		
Log Highways Investments * Log Redistribution * φ		-0.4516**	
		(0.19)	
Log Highways Investments * Log Equity * φ			0.1005
			(0.09)
Observations	5113	5115	5115
KP Wald F Statistic	58.859	62.838	72.128
Hansen J p-value	0.051	0.057	0.114
R^2	0.25	0.14	0.23

All regressions include the following set of control variables: GDP *per capita* in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

Table B5. Federal Highway Investments and Local Outcomes Growth by infrastructure project costs samples, 2007-2018: 2SLS IV Regressions

	1	2	3	4	5	6	7	8
	Cost Index 1 <= 50%	Cost Index 1 > 50%	Cost Index 2 <= 50%	Cost Index 2 > 50%	Cost Index 1 <= 50%	Cost Index 1 > 50%	Cost Index 2 <= 50%	Cost Index 2 > 50%
Log Highways Investments	0.0197**	0.0057	0.0182**	0.0071				
	(0.01)	(0.01)	(0.01)	(0.01)				
Log Highways Investments * φ					0.4769**	0.1326	0.4286**	0.1602
					(0.19)	(0.15)	(0.21)	(0.14)
Observations	2637	2478	2603	2512	2637	2478	2603	2512
KP Wald F Statistic	58.401	85.173	53.674	91.821	53.265	81.483	51.271	86.178
R^2	0.21	0.25	0.21	0.25	0.20	0.25	0.21	0.25

All regressions include the following set of control variables: GDP per capita in 2006; state fixed effects; municipality area; work force; agriculture share; exports share; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília; institutional quality; human capital. Robust standard errors reported in parentheses. * 0.1 ** 0.05 *** 0.01.

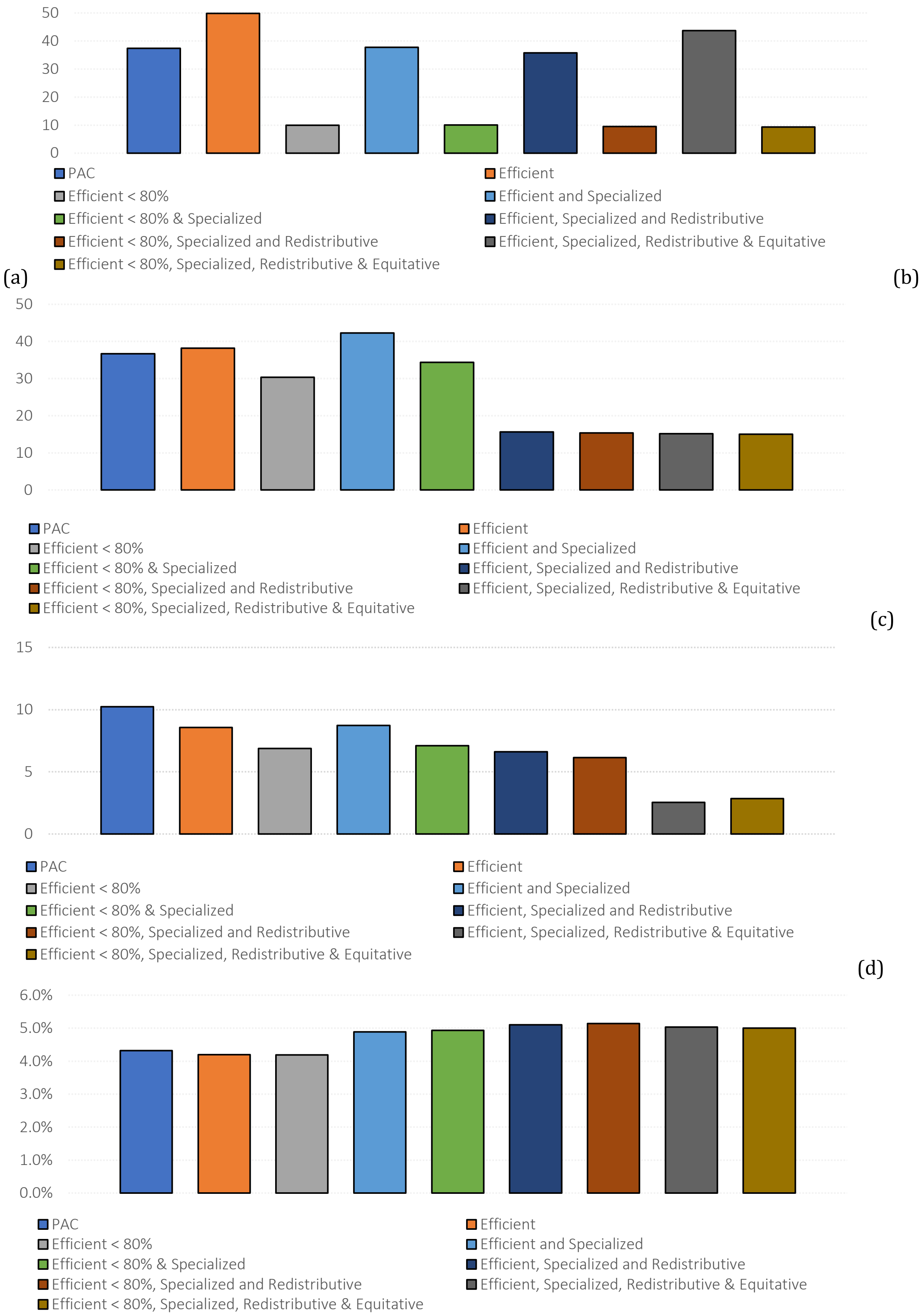
C. Section 6 Appendix

Table C2. Values of α by return rate

Return Rate	Condition (s)	$A^{sample}{}_{k}$	Source	Number of Municipalities
RR Average	Full sample	0.2928	Table 1 - Column 4	5,570
RR Efficient	Efficiency <= 50%	0.7068	Table 3 – Column 1	1,668
	50% < Efficiency & Efficiency <= 90%	0.6283	Table 3 – Column 1	
	Other	0.2928	Table 1 - Column 4	
RR Efficient & Road Specialized	50% < Efficiency & Efficiency <= 90% & 50% <Road Specialization	0.6897	Table 4 – Column 1 and Table 4 – Column 1	915
	Other	RR Efficient	-	
RR Efficient & Road Specialized & Redistributive & Equative	50% < Efficiency & Efficiency <= 90% & 50% < Road Specialization & 10% < Equity & Equity <= 50%	0.8161	Table 4 – Column 4	176
	50% < Efficiency & Efficiency <= 90% & 50% < Road Specialization & Equity <= 50% & Redistribution <= 50%	0.4452	Table 4 – Column 5	
	Other	RR Efficient & Road Specialized	-	
RR Efficient: 0 effect on efficiency upper 20%	80% < Efficiency	0	-	1,668
	Other	RR Efficient	-	
RR Efficient & Road Specialized: 0 effect on efficiency upper 20%	80% < Efficiency	0	-	915
	Other	RR Efficient & Road Specialized	-	
RR Efficient & Road Specialized & Redistributive & Equative: 0 effect on efficiency upper 20%	80% < Efficiency	0	-	176
	Other	RR Efficient & Road Specialized & Redistributive & Equative	-	

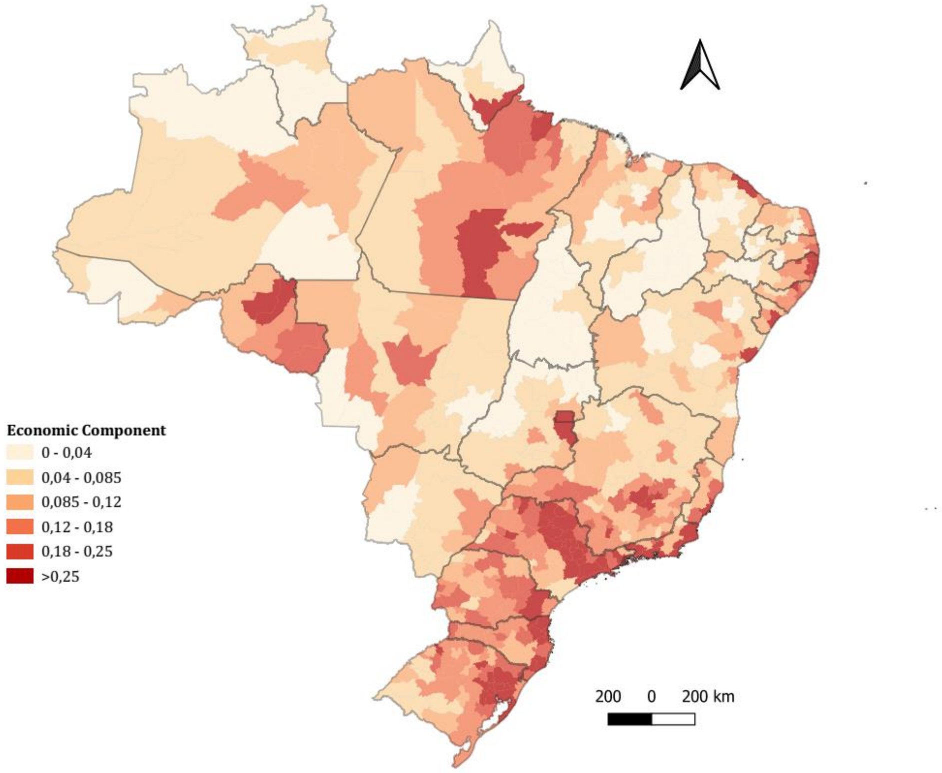
Source: authors' elaboration.

Figure CII.1. PAC and "Ideal Samples" averages: efficiency (a), redistribution (b), equity (c) and road specialization (d)

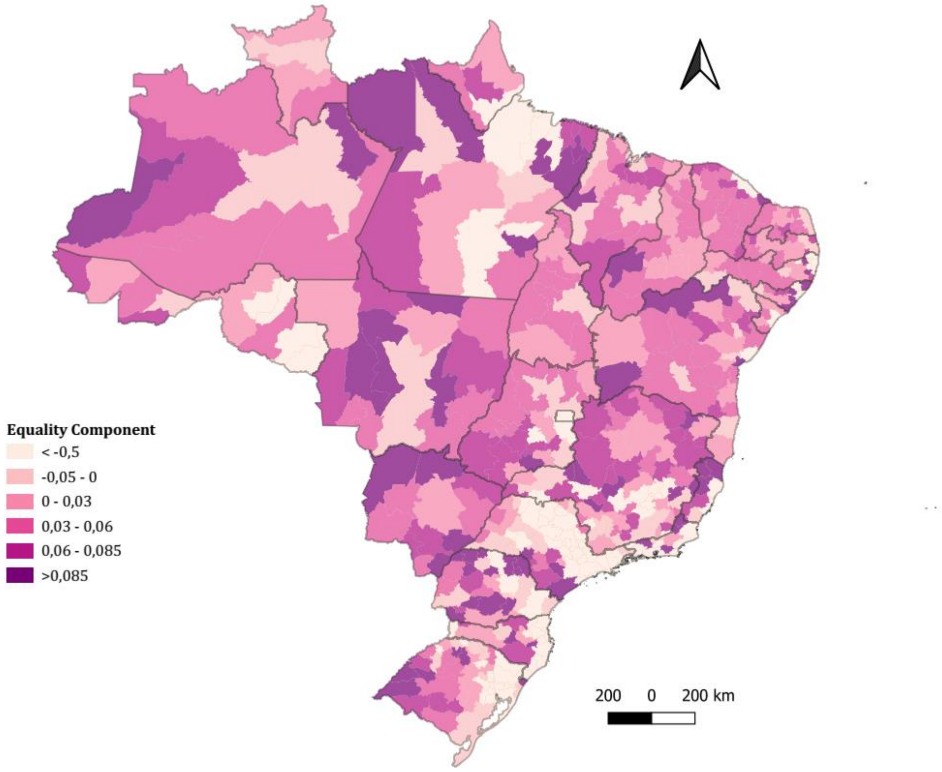


Source: authors' elaboration.

Figure C2. Return Rate to Highway Investments: Economic (a) and Equality (b) components



(a)



(b)

Source: authors' elaboration.

III. APENDIX - HIGHWAY INFRASTRUCTURE AND GREENHOUSE GAS EMISSIONS

A. Variables description and descriptive statistics

Table A1. Variables description

Type	Variable	Description	Source
Highway	Highway Investments	PAC Highway Investments (R$)	MINFRA
	Highway Intervention	1 if the municipality received PAC highway investments and 0 otherwise	MINFRA
	Highway Length	Road Length (km). Duplicates lanes were multiplied by 2, whilst single lanes were multiplied by 1	PNLT (2007) and MINFRA
IVs	Legal Protected Area	1 if the municipality is intersected by a legal protected area and 0 otherwise	MMA
	Environmental Embargos	1 if there was an environmental embargo in the municipality during this period and 0 otherwise	IBAMA
	Slope	Area with slope above 20% (which corresponds to hilly areas) (km^2)/Total area (km^2)	INPE
	Urban infrastructure	Building and infrastructure area (km^2)/ Total area (km^2)	MAPBIOMAS (Souza *et al.*, 2020)
	Population density	Population/Area (km^2)	IBGE
	Cost Index 1	First component of the MCA. Formula: $Cost\ Index\ 1 = 0.58 * Legal\ Protected\ Area + 0.51 * Embargos + 0.13 * Sloped\ Area + 0.08 * Urban\ Infrastructure.$	Medeiros *et al. (2024)*
	Cost Index 2	Second component of the MCA. Formula: $Cost\ Index\ 2 = 0.00 * Legal\ Protected\ Area + 0.01 * Embargos + 0.49 * Sloped\ Area + 0.58 * Urban\ Infrastructure$	Medeiros *et al. (2024)*
	LCP-MST Starting and Ending Road Points	Distance to the nearest LCP-MST hypothetical line using starting and ending road points as hubs	MINFRA
	Brasília Plan	Distance to the nearest Brasília Plan line (weighted by the municipality area share into the buffer zone)	Bird and Straub (2020)
	Potential Road Intervention Area	Distance to the nearest road segment classified as heavy traffic (D, E or F classification)	PNLT (2007)
	Non-Random Allocation Index	First component of the PCA. Formula: $Nonrandom\ Allocation\ Index = 0.58 * \text{LCP} - \text{MST Starting and Ending Road Points} + 0.58 * BrasíliaPlan - 0.57 * PotentialInterventionAreas$	Medeiros *et al. (2024)*
Dependent	CO2 emissions, 2018	CO2 emissions (ton) in 2018	SEEG
	CO2 emissions growth	Log(CO2 emissions, 2018) – Log(CO2 emissions, 2007)	SEEG
Moderators and Controls	CO2 emissions, 2007	CO2 emissions in 2007	SEEG
	GDP *per capita*, 2007	Gross Domestic Product (R$)/ Number of workers in 2007	IBGE
	GDP *per capita* squared, 2007	Gross Domestic Product (R$)/ Number of workers in 2007, squared	IBGE
	Population	Population in 2006	IBGE
	Population Density	Population in 2006/Geographic Area (km^2)	IBGE
	Road CO2 emissions (%)	Road sector CO2 emissions (ton)/ Total CO2 emissions (ton)	SEEG
	Capital-Labor ratio	Residential Capital Stock (R$)/Occupied Population	IBGE
	Exports share (%)	Municipal Exports (US$) / National Exports (US$)	MDIC
	Deforestation (1996-2006)	Variation in the deforestation rate between 1996 and 2006	MapBiomas
	Gini Index	Gini Index, 2000	IBGE
	Institutional Quality	Institutional quality municipal index (IQIM)	Ministry of Planning
	Human Capital (%)	Workers with graduate education/Total workers	RAIS/MTE

Distance to state road	Distance (km) to the nearest state road	MINFRA
Distance to railroad	Distance (km) to the nearest railroad	MINFRA
Distance to port	Distance (km) to the nearest federal port	MINFRA
Railways stations in 1920	Number of railways stations in 1920	Rede Ferroviária Federal S/A
Distance to Brasília	Distance (km) to the capital Brasília	IBGE

Source: authors' elaboration.

Table A2. Descriptive statistics

Variable	Mean	Std. Dev.
Highway Investments	1.270	3.447
Highway Intervention	0.125	0.331
Highway Length Growth (2007-2018)	0.375	0.985
Cost Index 1	-0.002	1.139
Cost Index 2	-0.003	1.013
Non-Random Allocation Index	3.000	1.635
LCP-MST Starting and Ending Road Points	1.956	2.119
Potential Road Intervention Area	2.288	2.516
Brasília Plan	2.455	2.749
CO2 Emissions Growth (2007-2018)	0.038	0.448
CO2 Emissions, 2007	11.591	1.310
GDP per capita, 2007	2.790	0.724
GDP per capita squared, 2007	8.307	4.395
Population	9.384	1.154
Population Density	3.131	1.414
Road CO2 emissions (%)	0.092	0.122
Capital-Labor ratio	2.899	0.574
Exports share (%)	0.000	0.001
Deforestation (1996-2006)	-0.006	0.216
Gini Index	-0.611	0.127
Institutional Quality	3.023	0.552
Human Capital (%)	0.001	0.004
Distance to state road	6.872	0.517
Distance to railroad	-0.348	1.623
Distance to port	5.462	0.946
Railways stations in 1920	3.200	1.924
Distance to Brasília	0.535	2.878

Source: authors' elaboration.

B. OLS estimates

Table B1. Federal Highway Investments and CO2 Emissions Growth (2007-2018): OLS Regressions

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Highways Investments	0.0042*	0.0051**	0.0142***	0.0078***	0.0061***	0.0057***	0.0056***	0.0056***	0.0055**	0.0056***	0.0056***	0.0057***	0.0057***	0.0057***	0.0064***	0.0065***	0.0065***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CO2 Emissions, 2007			-0.1164***	-0.2111***	-0.2246***	-0.2274***	-0.2273***	-0.2271***	-0.2242***	-0.2258***	-0.2258***	-0.2258***	-0.2232***	-0.2232***	-0.2292***	-0.2302***	-0.2303***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Population				0.1501***	0.1494***	0.1504***	0.1477***	0.1490***	0.1492***	0.1483***	0.1478***	0.1480***	0.1459***	0.1458***	0.1551***	0.1594***	0.1590***
				(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
GDP per capita					0.1058***	0.3258***	0.3027***	0.2971***	0.2894***	0.2920***	0.2896***	0.2903***	0.2909***	0.2909***	0.2892***	0.2947***	0.2943***
					(0.01)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
GDP per capita, square						-0.0339***	-0.0311***	-0.0297***	-0.0287***	-0.0290***	-0.0289***	-0.0290***	-0.0290***	-0.0290***	-0.0289***	-0.0292***	-0.0292***
						(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ratio Capital-Labor							0.0214	0.0224	0.0199	0.0225	0.0221	0.0221	0.0229	0.0229	0.0405**	0.0510***	0.0510***
							(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Exports by municipality (% total)								-7.9912	-8.2465	-8.1240	-8.2866	-8.2064	-8.3464	-8.3464	-2.8651	-2.6049	-2.7886
								(6.46)	(6.50)	(6.50)	(6.54)	(6.55)	(6.55)	(6.55)	(6.66)	(6.63)	(6.47)
Deforestation 1996-2006									0.1166***	0.1173***	0.1171***	0.1171***	0.1170***	0.1170***	0.1064***	0.1008***	0.1008***
									(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Gini Index										0.0569	0.0547	0.0548	0.0533	0.0533	0.0223	0.0090	0.0087
										(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Institutional Quality											0.0088	0.0090	0.0081	0.0081	0.0037	0.0028	0.0028
											(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Graduate education (% workers)												-0.6813	-0.5552	-0.5552	-0.7394	-0.5189	-0.5308
												(1.01)	(1.00)	(1.00)	(1.00)	(0.98)	(0.98)
Distance to Brasilia													0.0575***	0.0575***	0.1150***	0.1111***	0.1108***
													(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance to the nearest state road														0.0000	-0.0005	-0.0004	-0.0004
														(0.00)	(0.00)	(0.00)	(0.00)
Distance to the nearest port															0.0469***	0.0435***	0.0436***
															(0.01)	(0.01)	(0.01)
Distance to the nearest railroad																0.0152***	0.0153***
																(0.00)	(0.00)
Number of railway stations in 1920																	0.0005
																	(0.00)
Constant	0.0324***	-0.0180	1.5943***	1.4764***	1.3763***	1.0586***	1.0714***	1.0601***	1.0708***	1.1220***	1.1057***	1.1021***	0.6587***	0.6587***	-0.0884	-0.2179	-0.2121
	(0.01)	(0.06)	(0.11)	(0.10)	(0.10)	(0.13)	(0.13)	(0.13)	(0.13)	(0.14)	(0.14)	(0.14)	(0.22)	(0.22)	(0.25)	(0.25)	(0.25)
State FE	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	5559	5559	5559	5553	5553	5553	5496	5496	5494	5494	5490	5490	5489	5489	5489	5489	5489
R^2	0.001	0.076	0.140	0.210	0.223	0.225	0.228	0.228	0.231	0.231	0.231	0.231	0.232	0.232	0.235	0.238	0.238

* p<0.1, ** p<0.05, *** p<0.01

Table B2. Federal Highway Investments and CO2 Emissions Growth (2007-2018): OLS Regressions

	1 All	2 Roads	3 Energy	4 Land Use	5 Agriculture
Highways Investments	0.0065***	0.0212***	0.0220***	-0.0043	0.0035**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CO2 Emissions, 2007	-0.2303***	-0.6416***	-0.6397***	-0.2784***	-0.0122**
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Population	0.1590***	0.7530***	0.7384***	0.1171***	-0.0122*
	(0.01)	(0.03)	(0.03)	(0.02)	(0.01)
GDP per capita	0.2943***	0.7858***	1.0141***	0.1962	0.0917
	(0.06)	(0.17)	(0.15)	(0.13)	(0.06)
GDP per capita, square	-0.0292***	-0.0679***	-0.0824***	-0.0347*	-0.0128
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Ratio Capital-Labor	0.0510***	0.0267	0.0574	-0.0428	0.0028
	(0.02)	(0.06)	(0.06)	(0.04)	(0.02)
Exports by municipality (% total)	-2.7886	-19.1449**	-9.8519	-2.7054	-4.5264
	(6.47)	(8.43)	(10.89)	(15.78)	(5.56)
Deforestation 1996-2006	0.1008***	0.0233	0.0379	0.1157	-0.0305
	(0.03)	(0.08)	(0.08)	(0.08)	(0.03)
Gini Index	0.0087	0.1316	0.1251	0.3269***	0.0054
	(0.05)	(0.14)	(0.14)	(0.10)	(0.04)
Institutional Quality	0.0028	-0.0464	-0.0553	-0.0314	0.0210*
	(0.01)	(0.04)	(0.03)	(0.03)	(0.01)
Graduate education (% workers)	-0.5308	-2.1816	-2.6223	-3.2895	0.3049
	(0.98)	(2.64)	(2.58)	(2.22)	(0.94)
Distance to Brasilia	0.1108***	-0.1451**	-0.1663***	0.3517***	0.0559**
	(0.02)	(0.06)	(0.06)	(0.05)	(0.02)
Distance to the nearest state road	-0.0004	-0.0013	0.0093	0.0012	-0.0048
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
Distance to the nearest port	0.0436***	0.0768***	0.0536***	0.0887***	0.0599***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Distance to the nearest railroad	0.0153***	0.0125	0.0096	0.0317***	0.0037
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
Number of railway stations in 1920	0.0005	-0.0130***	-0.0115***	0.0086*	0.0051*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	-0.2121	-2.0948***	-2.1991***	-0.7282	-0.5189**
	(0.25)	(0.65)	(0.62)	(0.54)	(0.24)
State FE	Y	Y	Y	Y	Y
Observations	5489	5489	5489	5489	5489
R^2	0.238	0.551	0.552	0.238	0.156

* p<0.1, ** p<0.05, *** p<0.01

C. Heterogeneous road impacts on CO emissions growth

Table C1. Federal Highway Investments and CO2 Emissions Growth (2007-2018) - Elasticity (α*φ): Heterogeneous Point Impacts

Sample	Roads CO2 Emissions (%)			Log GDP per capita			Log Ratio Capital-Labor			Population Density			Deforestation (1996-2006)			Population		
	Coeff.	[95% Conf. Interval]		Coeff.	[95% Conf. Interval]		Coeff.	[95% Conf. Interval]		Coeff.	[95% Conf. Interval]		Coeff.	[95% Conf. Interval]		Coeff.	[95% Conf. Interval]	
10%	0.028	0.012	0.045	0.035	0.010	0.061	0.040	0.017	0.064	0.022	0.005	0.039	0.017	0.001	0.032	0.034	0.013	0.055
25%	0.027	0.011	0.043	0.031	0.011	0.052	0.031	0.014	0.049	0.018	0.004	0.033	0.021	0.006	0.036	0.030	0.012	0.048
50%	0.024	0.009	0.040	0.023	0.009	0.038	0.020	0.006	0.034	0.015	0.001	0.029	0.024	0.010	0.039	0.024	0.010	0.039
75%	0.018	0.003	0.033	0.016	0.002	0.031	0.010	-0.006	0.026	0.013	-0.002	0.027	0.027	0.013	0.042	0.019	0.004	0.033
90%	0.004	-0.017	0.026	0.010	-0.010	0.030	0.003	-0.018	0.023	0.009	-0.008	0.026	0.032	0.016	0.048	0.012	-0.005	0.029

All regressions include the following set of control variables: CO2 emissions in 2007; state fixed effects; population; GDP *per capita*; GDP *per capita*, square; capital-labor ratio; exports share; 1996-2006 deforestation; Gini index; institutional quality; human capital; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília. Robust standard errors reported in parentheses.

D. Robustness checks

Table D1. Federal Highway Investments and CO2 Emissions Growth (2007-2018): 2SLS IV Regressions

	1	2	3	4	5	6	7	8	9	10
Second stage	All	Roads	Energy	Land Use	Agriculture	All	Roads	Energy	Land Use	Agriculture
Highway Intervention	0.2457***	1.3293***	1.1520***	0.5282***	-0.0300					
	(0.08)	(0.22)	(0.21)	(0.15)	(0.06)					
Highway Length Growth						0.1010***	0.5500***	0.4765***	0.2168***	-0.0125
						(0.03)	(0.09)	(0.09)	(0.06)	(0.02)
First stage										
Non-Random Allocation Index	-0.0476***	-0.0483***	-0.0485***	-0.0479***	-0.0484***	-0.1158***	-0.1167***	-0.1172***	-0.1168***	-0.1158***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	5142	5142	5142	5142	5142	5142	5142	5142	5142	5142
KP Wald F Statistic	347.700	357.426	358.430	353.998	356.778	254.337	257.484	258.945	257.720	249.632
R^2	0.23	0.51	0.53	0.22	0.15	0.22	0.49	0.51	0.21	0.15

All regressions include the following set of control variables: CO2 emissions in 2007; state fixed effects; population; GDP *per capita*; GDP *per capita*, square; capital-labor ratio; exports share; 1996-2006 deforestation; Gini index; institutional quality; human capital; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília. Robust standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table D2. Federal Highway Investments and CO2 Emissions Growth (2007-2018): 2SLS IV Regressions

	1	2	3	4	5	6	7
Second stage							
Log Highways Investments	0.6151***	0.5317***	0.5891***	0.6112***	0.5994***	0.6108***	0.6430***
	(0.17)	(0.19)	(0.20)	(0.18)	(0.17)	(0.18)	(0.17)
First stage							
Cost Index 1	0.0069***				0.0074***	0.0065***	0.0066***
	(0.00)				(0.00)	(0.00)	(0.00)
Cost Index 2	0.0113***				0.0104***	0.0122***	0.0111***
	(0.00)				(0.00)	(0.00)	(0.00)
Non-Random Allocation Index	-0.0204***						
	(0.00)						
LCP-MST Starting and Ending Road Points		-0.0150***			-0.0151***		
		(0.00)			(0.00)		
Potential Road Intervention Area			0.0128***			0.0130***	
			(0.00)			(0.00)	
Brasília Plan				-0.0121***			-0.0121***
				(0.00)			(0.00)
Observations	5131	5142	5142	5142	5131	5131	5131
KP Wald F Statistic	116.199	310.790	288.358	326.464	108.806	102.897	112.932
R^2	0.22	0.23	0.22	0.22	0.22	0.22	0.22

All regressions include the following set of control variables: CO2 emissions in 2007; state fixed effects; population; GDP *per capita*; GDP *per capita*, square; capital-labor ratio; exports share; 1996-2006 deforestation; Gini index; institutional quality; human capital; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília. Robust standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table D3. Federal Highway Investments and CO2 Emissions Levels (2018): 2SLS IV Regressions

	1	2	3	4	5
Second stage	All	Roads	Energy	Land Use	Agriculture
Log Highways Investments * φ	0.5770***	3.0916***	2.6802***	1.2362***	-0.0702
	(0.18)	(0.51)	(0.49)	(0.36)	(0.14)
First stage					
Non-Random Allocation Index	-0.0203***	-0.0208***	-0.0208***	-0.0205***	-0.0207***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	5142	5142	5142	5142	5142
KP Wald F Statistic	332.469	343.528	344.349	338.341	340.560
R^2	0.90	0.67	0.71	0.77	0.95

All regressions include the following set of control variables: CO2 emissions in 2007; state fixed effects; population; GDP *per capita*; GDP *per capita*, square; capital-labor ratio; exports share; 1996-2006 deforestation; Gini index; institutional quality; human capital; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília. Robust standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table D4. Federal Highway Investments and CO2 Emissions Growth (2007-2018): 2SLS IV Regressions

	1	2	3	4
Second stage	All – Excluding municipalities in Amazon	All – Excluding municipalities in Pará	All – Excluding municipalities in São Paulo	Land Use – Excluding municipalities in Amazon
Log Highways Investments * φ	0.6137***	0.6938***	0.4980***	1.2692***
	(0.20)	(0.19)	(0.19)	(0.44)
First stage				
Non-Random Allocation Index	-0.0187***	-0.0196***	-0.0208***	-0.0188***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	4383	5004	4585	4383
KP Wald F Statistic	243.015	311.298	301.298	246.454
R^2	0.20	0.20	0.24	0.22

All regressions include the following set of control variables: CO2 emissions in 2007; state fixed effects; population; GDP *per capita*; GDP *per capita*, square; capital-labor ratio; exports share; 1996-2006 deforestation; Gini index; institutional quality; human capital; distance to the nearest state road; distance to nearest railroad; distance to nearest port; railways stations in 1920; distance to Brasília. Robust standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

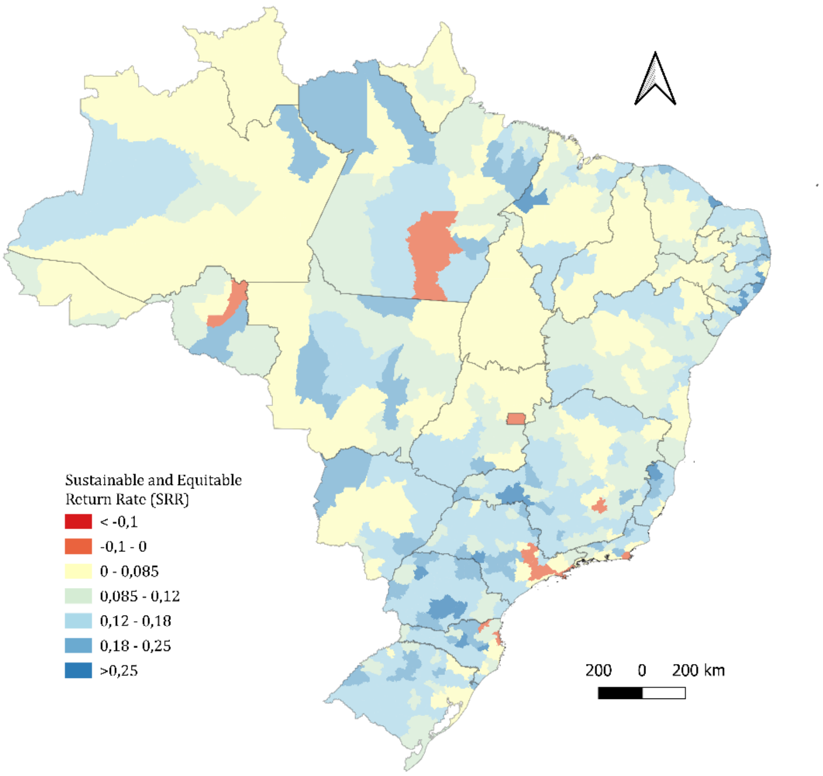
E. Sensitivity analysis: land use change CO2 emissions

Table E1. Sensitivity Analysis: Drops in Land Use Change CO2 emissions

Drop in Land Use Change CO2 Emissions	Average ERR	Average SERR
25%	2.54%	17.46%
50%	2.17%	17.63%
75%	1.79%	18.21%
100%	1.42%	18.58%

Source: authors' elaboration.

Figure E1. Sustainable Return Rates to Highway Investments: SERR

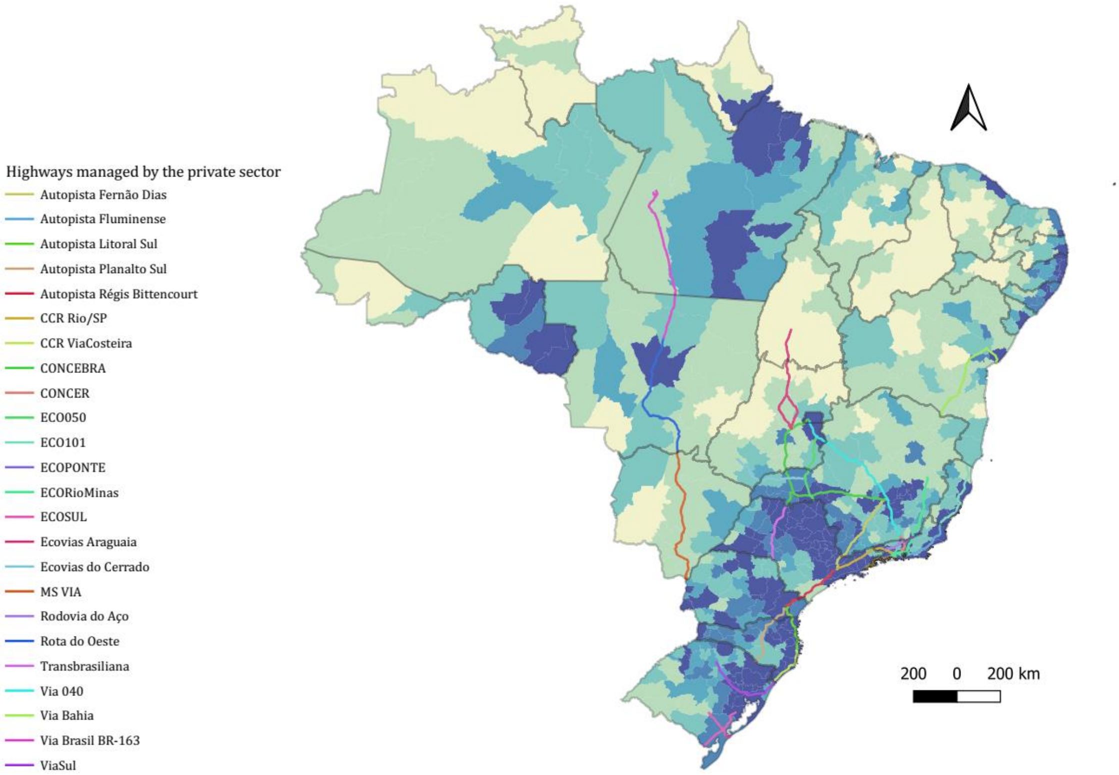


Source: authors' elaboration.

IV. APENDIX - BRINGING HIGHWAY INVESTMENTS MORE EFFICIENT, REDISTRIBUTIVE, AND SUSTAINABLE

A. SERR and road features

Figure A1. Privately managed highways and SERR



Source: authors' elaboration.

Table A1. SERR and its components: public and private management

Variable	Public	Private
EC	0.0970	0.1701
SC	0.0120	-0.0415
GEC	0.0123	0.0143
SERR	0.0965	0.1143

Source: authors' elaboration.

Figure A2. Correlation matrix: SERR and road features

	EC	SC	GEC	SERR	Efficiency (GDP/Road Stock)	Redistribution (GDP *per capita)*	Equity (Road Stock/Area)	Road specialization (ϕ)
EC	1.000							
SC	-0.870	1.000						
GEC	0.179	-0.075	1.000					
SERR	-0.050	0.534	0.069	1.000				
Efficiency (GDP/Road Stock)	0.400	-0.420	0.066	-0.166	1.000			
Redistribution (GDP per capita)	0.556	-0.431	0.248	0.066	0.198	1.000		
Equity (Road Stock/Area)	0.552	-0.478	-0.385	0.021	0.284	0.326	1.000	
Road specialization (ϕ)	0.362	-0.079	0.383	0.435	0.036	0.291	0.095	1.000

B. State level SERR components

Figure B1. Economic Component (EC): state-level

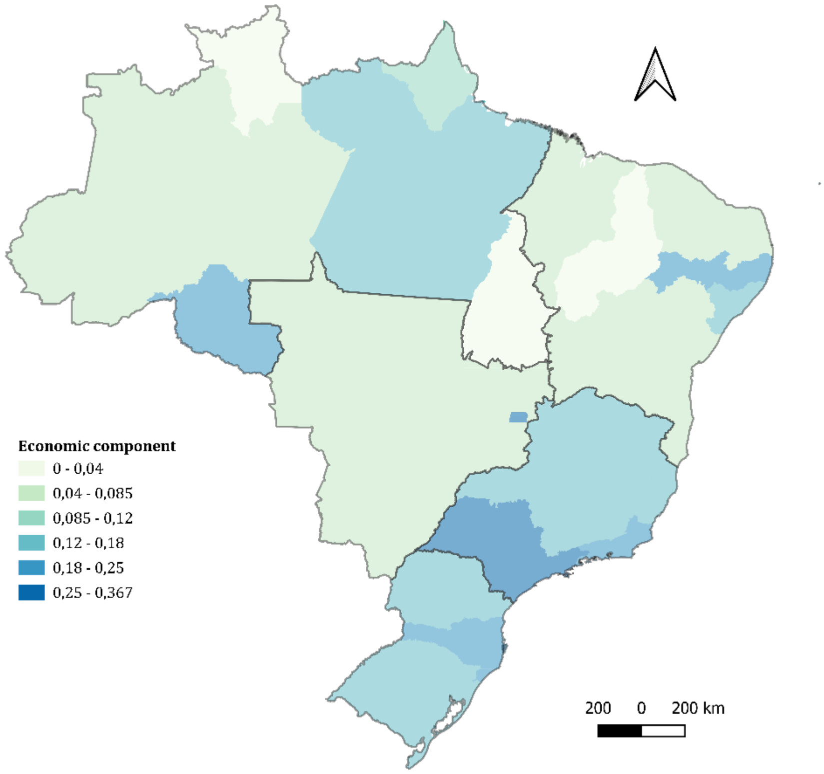


Figure B2. Social Component (SC): state-level

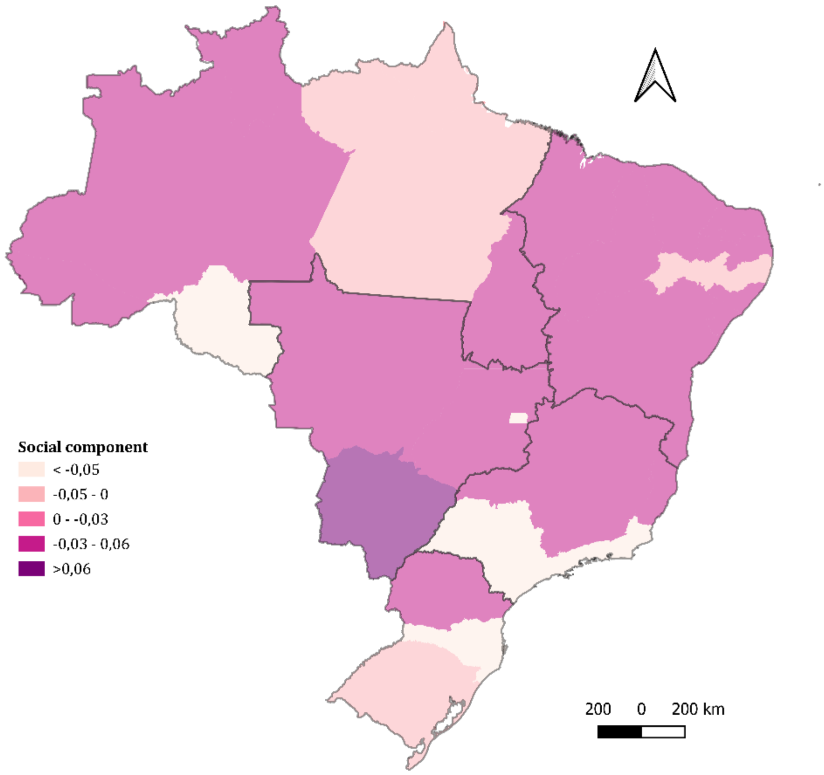
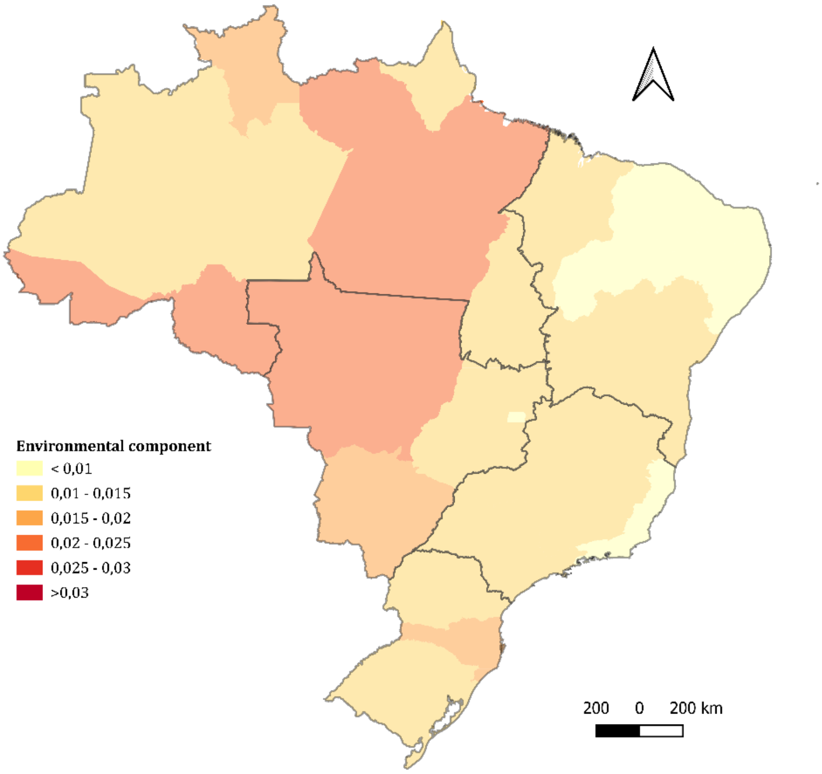


Figure B3. Environmental component (GEC): state level



C. SKATER application

We start by creating 27 clusters as we have the state administrative boundaries as a benchmark for comparison. The results in terms of the total within-cluster sum of squares, the between-cluster sum of squares, and the ratio of between to total sum of squares as well as a detailed description of the SKATER application results can be seen in Table C1. Whilst the resulted clusters fitted well following the sum of squares statistics, some clusters are very small, and the practical application of those regions is doubt. Then, we generate 27 clusters again but constraining then to have at least 1% and 2% of the national geographical area. The cluster seems to be more suitable to represent highway policy zones, but some small regions remain.

Next, we compute clusters only based on the geographic area constraint. We force each cluster geographic area to be from 1 to 5 percent of the national geographic area. In those cases, we do not impose any constraint in terms of the number of clusters. Then, the resulted SKATER contains 64, 34, 20, 17 and 12 zones for the 1, 2, 3, 4 and 5 percent constraints, respectively. We can see a huge drop in the ratio of between to total sum of squares from the 1% to the 2% geographic constraint, a small from the 2% to the 3% geographic constraint, and again a higher decrease from the 3% to 4% geographic constraint. This result suggests that using the 1% or 3% geographic constraints appears to be more suitable. In addition, we can observe an apparently more appropriate highway zone size when taking geographic constraints higher than 3%.

Thereafter, we generate new clusters by imposing the number of regions equal to 10 and varying the clusters geographic area to be from 1% to 5% of the national graphic area. Obviously, the ratio of between to total sum of squares decreases in comparison with the previous SKATER applications as we impose a reduced number of clusters to be computed. On the other hand, the created highway zones seem to suitably characterize the regions in terms of their economic, social, and environmental features. We observe small drops in the ratio of between to total sum of squares from the 1% to the 3% geographic constraint, then a considerable decrease is observed from the 3% to the 4% geographic constraint. This result suggests choosing the clusters generated following the 3% geographic constraint.

As robustness checks, we try several combinations varying the number of clusters from 11 to 19 or 20, and taking the geographic constrains of 1% and 3%. Then, 19 new combinations are generated. As expected, even varying the number of clusters, the results from the 1% geographic constraint appears to have practical issues, as very small zones remain. Regarding the 3% geographic constraint combinations, the increasing in the number of clusters rises the ratio of between to total sum of squares very slightly, then suggesting that we might hold the 10 clusters generated using the same geographic constraint.

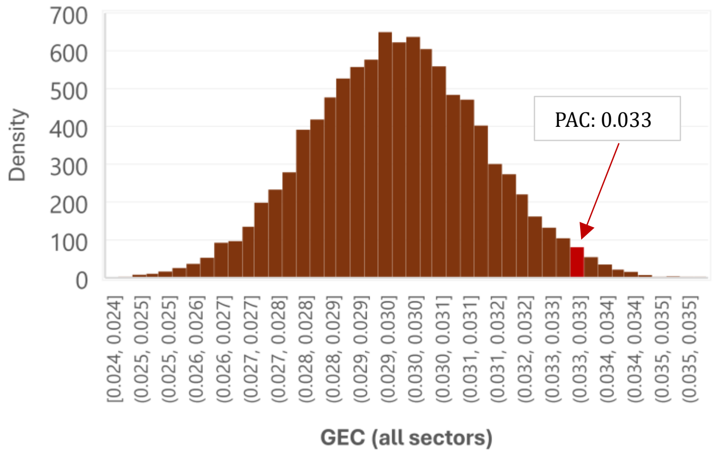
Table C1. SKATER application results

Number of regions	Minimum Geographic Area (%)	The total within-cluster sum of squares	The between-cluster sum of squares	The ratio of between to total sum of squares
27	-	423.686	1103.310	0.723
27	2	854.201	672.799	0.441
27	1	714.392	812.608	0.532
64	1	633.789	893.211	0.585
34	2	848.798	678.202	0.444
20	3	915.733	611.267	0.400
17	4	1023.170	503.831	0.330
12	5	1044.200	482.804	0.316
10	5	1057.960	469.038	0.307
10	4	1059.090	467.913	0.306
10	3	961.956	565.044	0.370
10	2	957.929	569.071	0.373
10	1	923.023	603.977	0.396
11	3	952.034	574.966	0.377
12	3	944.264	582.736	0.382
13	3	938.892	588.108	0.385
14	3	933.831	593.169	0.388
15	3	925.125	601.875	0.394
16	3	922.531	604.469	0.396
17	3	920.176	606.824	0.397
18	3	917.854	609.146	0.399
19	3	916.512	610.488	0.400
11	1	906.461	620.539	0.406
12	1	890.722	636.278	0.417
13	1	875.721	651.279	0.427
14	1	832.947	694.053	0.455
15	1	818.659	708.341	0.464
16	1	804.906	722.094	0.473
17	1	791.255	735.745	0.482
18	1	779.233	747.767	0.490
19	1	768.133	758.867	0.497
20	1	760.835	766.165	0.502

In short, we get two main and conflicting (and expected) results from the SKATER applications. First, as we increase the number of clusters to around 27 following the state boundaries as a benchmark, the results in terms of the ratio of between to total sum of squares improve. Nonetheless, when we increase the number of clusters, very small clusters are generated, especially when the geographical area restrictions are zero or smaller. The result suggests that the practical use of those zone for public policy aims might be hampered. Second, by limiting the number of clusters to be around 10, the zones generated seem to appropriate represent regional characteristics for road policy development purposes. However, the ratio of between to total sum of squares is smaller in comparison with the 27 clusters. That said, we opt for the 10 clusters constrained by the 3% geographic area requirement, which seems to represent the Brazilian reality in a more reasonable way. As a robustness check, we also use the 27 clusters restricted by 1% of the national geographic area. Due to the hierarchical nature of the SKATER technic, the 27 clusters complement the 10 clusters as the latter are partitions of the former. Then, if the generated 10 highway policy zones are considered too large, the 27 clusters results provide mechanisms to visualize and evaluate more disaggregated road policy zones. The preferred clusters generated applying the SKATER can be seen in Figure 1.

D. Simulation results: sensitivity analysis

Figure D1. Simulation results: EC, SC, GEC and SERR considering all sectors in GHG emissions



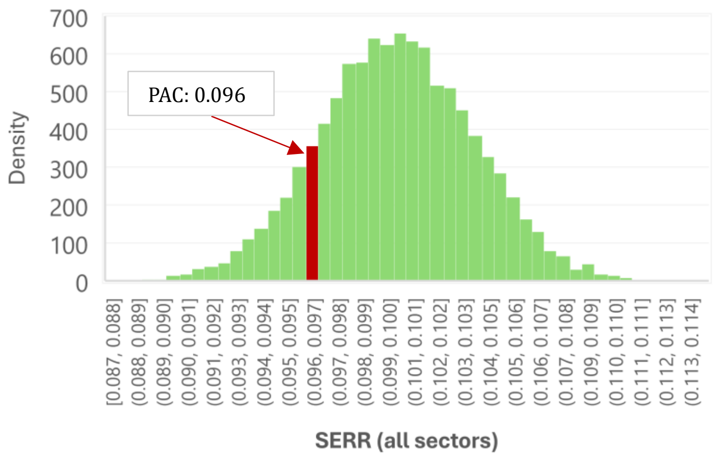


Table D1. Summary of results: simulations using state and highway policy zone levels

Sample	EC	SC	GEC	SERR
State Level				
PAC	0.1503	-0.0169	0.0137	0.1198
Simulations	0.1281	-0.0109	0.0124	0.1048
Highway Policy Zone Level				
PAC	0.1477	-0.0177	0.0134	0.1166
Simulations	0.1358	-0.0111	0.0141	0.1105