Unlocking Urban Mobility: Impact Evaluation of Metro Line 1 on Travel Time and Transportation Costs in Metropolitan Lima and Callao

Edson Huamaní* Tania Paredes†

August 12, 2026

Abstract

This study estimates the causal effect of Line 1 of the Basic Metro Network of Lima and Callao (RBMLC) on the travel time and transportation costs of its users. Exploiting the quasi-experimental variation generated by the spatial configuration of the stations, the empirical strategy implements an Instrumental Variables estimator with High-Dimensional Fixed Effects (IV-HDFE) that corrects for the self-selection bias in mode choice.

The results indicate that using Line 1 reduces in-vehicle time by 27.80 minutes and increases the effective travel speed by 11.8 km/h relative to the counterfactual of substitute transport. On the economic dimension, we estimate a reduction of S/ 1.23 in the monetary cost per trip, an effect that is progressively concentrated among users of lower socioeconomic status (S/ 2.03) and among trips during the evening peak (S/ 2.39), positioning Line 1 as an implicit instrument of income redistribution. These benefits, however, entail a logistical penalty: the system increases waiting times by 2.71 minutes on average, an effect that rises to 10.9 minutes during the evening peak, reflecting demand saturation at the busiest stations. These findings validate the social profitability of the project and identify concrete operational frictions whose resolution is a priority for maximizing the benefits of future network expansions.

A preliminary quantification suggests that the in-vehicle time savings alone imply a capital-cost recovery period of between four and seven years, which supports the social profitability of the project.

Keywords: Peru, Urban Transport, Metro Systems, Infrastructure Evaluation
JEL Classification: R40, R42, O18

*Corresponding author. Pontificia Universidad Católica del Perú and Universidad Peruana de Ciencias Aplicadas; e-mail: ehuamani@pucp.pe, ORCID: https://orcid.org/0000-0002-3611-1891

†Pontificia Universidad Católica del Perú and Universidad Peruana de Ciencias Aplicadas; e-mail: tania.paredes@pucp.edu.pe, ORCID: https://orcid.org/0000-0003-0190-9753

1 Introduction

Urban transport is a central determinant of the productivity, welfare, and environmental sustainability of modern economies, as it structures households' decisions on residence, employment, and consumption. In Peru, however, the transport system faces structural challenges that limit its functionality, sustainability, and capacity to respond to a growing and increasingly complex demand. These limitations are particularly acute in Metropolitan Lima and Callao, where high population density, a lagging road infrastructure, and the absence of an integrated mass transit system converge. The outcome is not confined to operational inefficiencies and a sustained deterioration in service quality; it also manifests in negative economic, social, and environmental externalities that erode the city's competitiveness and its residents' quality of life.

The magnitude of the problem places Lima among the most critical cases worldwide. According to the *TomTom Traffic Index*, in 2024 the city ranked first among 501 evaluated cities in terms of time lost during peak hours, with an average of 155 hours per driver per year (TomTom Traffic Index (TTI), 2024). Far from being a transitory phenomenon, congestion describes a worsening trend over time: according to the National Time-Use Survey (ENUT), the average commuting time to work rose from 0.84 to 1.33 hours per day between 2010 and 2024, a 57 % increase (Céspedes-Reynaga, 2025). In urban areas, workers spend on average 1 hour and 21 minutes per day commuting, about 11 minutes more than their rural counterparts, a gap that reflects the growing pressure on metropolitan mobility systems (Instituto Nacional de Estadística e Informática, 2025).

The consequences of this congestion extend beyond individual time losses and acquire a first-order macroeconomic dimension. The Central Reserve Bank of Peru estimates that the annual cost of vehicular congestion in Lima amounts to approximately 2.4 % of GDP, arising both from excess fuel consumption during peak hours (0.4 % of GDP) and from the economic value of the working hours lost in commuting (2.0 % of GDP) (Banco Central de Reserva del Perú, 2024). These figures are consistent with the international evidence for Latin America, which places congestion losses between 0.5 % and 1.1 % of GDP, and confirm that travel time operates as an implicit tax on labor productivity and household welfare.

Against this backdrop, the expansion and consolidation of efficient and environmentally sustainable mass transit systems, such as Line 1 of the Basic Metro Network of Lima and Callao (RBMLC), emerges as a central component of the transformation of urban mobility toward a more integrated and resilient model. Rigorously evaluating the effects of this type of intervention is therefore indispensable for guiding investment and expansion decisions for the city's transport infrastructure. This need is embedded, moreover, in a consolidated international literature that has documented the multidimensional benefits of metro systems in urban settings.

Indeed, prior evidence highlights the contribution of these systems to the reduction of travel times (Oviedo et al., 2019; Kim et al., 2021), the alleviation of vehicular congestion (Hu et al., 2024), and improvements in intra-urban accessibility, factors that drive local economic development (Cheng and Chen, 2021; Välilä, 2025). These systems also shape processes of urban transformation, through land-value appreciation and gentrification (Bocarejo et al., 2014; Scholl et al., 2018; Wu et al., 2020; Contreras et al., 2024), while helping to mitigate various transport externalities (Morales, 2024). Recent studies additionally emphasize the qualitative value of rail travel: the design of metro systems allows users to undertake productive or recreational activities during their trips, which improves the travel experience and influences mode choice (de Sá et al., 2025).

In the Peruvian case, the evidence shows that interventions such as the Metropolitano and Line 1 of the Lima Metro have generated significant socioeconomic and environmental effects in Metropolitan Lima and Callao (Scholl et al., 2018; Oviedo et al., 2019; Velásquez, 2023; Alba Vivar, 2024; Contreras et al., 2024; Morales, 2024). The sustainability of these benefits over time, however, depends on robust asset-management frameworks that ensure efficient maintenance, renewal, and expansion of the infrastructure (Mohammadi et al., 2019), as well as on planning mechanisms for service disruptions, such as rapid-response protocols and coordination with alternative transport systems (for example, replacement buses), that preserve operational resilience and user confidence (Zhang et al., 2021).

The relevance of these interventions in developing economies, however, is not free of debate. Some authors question the fiscal viability of highly segregated metro systems and propose lower-cost alternatives, such as light rail transit systems (Allport, 1986). Others argue that the expansion of metro systems remains a valid strategy in high-density urban corridors, provided that it is supported by sound financial planning and by institutional frameworks oriented toward efficiency, environmental sustainability, and territorial equity (Giannopoulos, 1981). This debate makes rigorous empirical evidence on the actual returns to investment especially valuable, precisely the type of evidence that remains scarce for cities in emerging economies.

The present study seeks to contribute to closing that gap. Its objective is to evaluate the causal impact of Line 1 of the RBMLC on key urban-mobility outcomes, in particular: (i) travel times and (ii) individual transportation costs. As a complement, we examine its effects on operational performance indicators, such as effective speed and waiting times. Given the potential problem of non-random selection into system use, the empirical strategy adopts an instrumental-variables approach that identifies the causal effects of mass transit infrastructure on these outcomes. In doing so, the research provides empirical evidence on the socioeconomic returns to mass transit systems in cities of emerging economies, and offers relevant inputs for the design and expansion of sustainable urban-mobility policies.

The evaluation draws on information from the representative surveys of the Urban Mobility Plan for Lima and Callao (PMU) 2023 (ATU, 2023), and on technical documents of the ATU (ATU, 2024, 2025). From these sources, we specify an instrumental-variables model to estimate the effects of Line 1 on travel times and costs.

Beyond the Peruvian case, this work offers two lessons of general relevance for the design of mass transit systems in cities of emerging economies. First, it documents how a subsidized flat fare operates as an implicit redistributive mechanism, whose progressivity arises from the interaction between the fare structure and income-differentiated travel patterns. Second, it shows that the capture of these benefits is bounded by the "last-mile" access constraint, which is particularly binding for lower-income households. As a complement, the magnitude of the estimated time savings allows a valuation of the social return to the investment, a direct input for network-expansion decisions.

The document is organized as follows. Section 2 describes the intervention, detailing its objectives, implementation, and main features. Section 3 develops the review of the literature on urban transport systems, with particular attention to the evidence in developing economies. Section 4 sets out the theory of change underlying the intervention, identifying the problems it addresses, the mechanisms through which it operates, and the expected outcomes. Section 5 presents descriptive statistics on the sample and the context of the intervention. Section 6 details the empirical strategy and its identification assumptions. Section 7 reports the results, and Section 8 concludes with policy recommendations aimed at improving the effectiveness of future interventions.

2 The Basic Metro Network of Lima and Callao (RBMLC)

The Basic Metro Network of Lima and Callao (RBMLC) is an Electric Mass Transit System designed to provide adequate transport infrastructure to the high-density zones and peripheral areas of Lima and Callao, with the aim of enabling efficient mobility in terms of both cost and travel time (Supreme Decree No. 059-2010-MTC; Supreme Decree No. 001-86-MIPRE). Since its initial conception, the system has undergone multiple revisions: its original design comprised six metro lines integrating different districts of Lima and Callao (Llosa and Panizza, 2015) (Supreme Decree No. 059-2010-MTC).

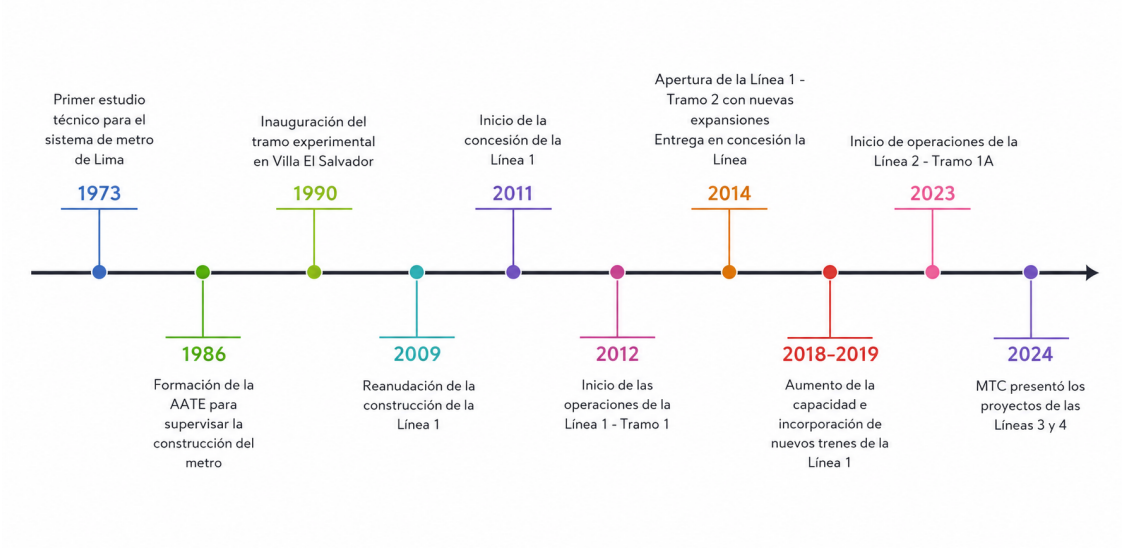
The formulation of the RBMLC rests on a public-policy framework that recognizes urban mobility as a right enabling economic, social, and environmental development. Through Supreme Decree No. 001-86-MIPRE, subsequently elevated to the rank of law via Law No. 24565, the Peruvian State declared the implementation of an Electric Mass Transit System to be of public necessity and of preferential social interest, thereby establishing the legal and policy bases for a structural transformation of transport in Metropolitan Lima and Callao.

Within this framework, the RBMLC is conceived as an intervention aimed primarily at increasing the efficiency, accessibility, and sustainability of the urban transport system through a high-capacity electric rail network. The intervention seeks to address the challenges of urban transport along four dimensions (GyM Ferrovías, 2014): (1) chronic structural congestion, reflected in the saturation of the main metropolitan corridors and in the associated productivity losses; (2) the fragmentation and informality of traditional transport services, characterized by an atomized supply, overlapping routes, and the absence of operational planning; (3) the intensification of the environmental footprint arising from dependence on motorized vehicles and their associated greenhouse-gas emissions; and (4) deficits in functional connectivity, whereby limited coverage and weak intermodal integration restrict equitable access to essential services, labor markets, and networks of urban opportunity. Accordingly, the RBMLC is structured around a long-term results framework that links strategic inputs with outputs, intermediate outcomes, and final impacts.

Regarding its initial implementation, the Autonomous Authority of the Electric Mass Transit System of Lima and Callao (AATE) was created in 1986, responsible for supervising the construction of Line 1 of the Lima Metro. In 1990, a 1.5-kilometer pilot segment was inaugurated and, over the following decade, the main civil works and system facilities were developed, along with the incorporation of 32 additional electric cars. Nonetheless, despite this progress in infrastructure and equipment, the project did not become fully operational during that period (Corporación Andina de Fomento (CAF), 2015).

The project resumed in 2009, when the Construction Contract was signed between the AATE, Provías Nacional, and the Lima Electric Train Consortium (CTEL), formed by Odebrecht and GyM S.A. Following the approval of financing by the CAF, the CTEL restarted the implementation works of Section 1 in 2010 and of Section 2 in 2011 of Line 1 of the RBMLC Comisión Investigadora del Caso Lava Jato, 2018. The total investment in the implementation of Line 1 amounted to US$ 671.7 million, according to the 2012 concession contract (OSITRAN, 2025).

Figure 1: Timeline of the main milestones in the formulation, construction, and operation of the Basic Metro Network of Lima and Callao



Source: Authors' own elaboration based on Corporación Andina de Fomento (CAF) (2015) and Comisión Investigadora del Caso Lava Jato (2018).

Line 1 of the RBMLC began operations in two stages: the first in 2012 (22.1 kilometers) and the second in 2014 (11.9 kilometers). Subsequently, its transport capacity was expanded by increasing the number of trains and cars during 2018 and 2019 (GyM Ferrovías, 2024). Since 2019, Line 1 has connected 11 districts of Metropolitan Lima along a 34-kilometer corridor, operated with 44 trains and 26 stations (GyM Ferrovías, 2018).

The operation and maintenance of Line 1 are managed under a co-financed concession model, which has maintained a socially regulated fare since the start of operations in 2012 (OSITRAN, 2024). Service frequency varies between 3.5 and 6 minutes, depending on operating conditions.

Figure 2: Basic Metro Network of Lima and Callao (RBMLC) – Line 1



Source: Authors' own elaboration based on Corporación Andina de Fomento (CAF) (2015) and Comisión Investigadora del Caso Lava Jato (2018).

Regarding the progress of the remaining lines, in December 2014 the concession for the co-financed construction of Line 2 of the Lima and Callao Metro was awarded, a 27-kilometer project with an additional 8-kilometer branch, whose first operational segment came into service in 2023 (Ministerio de Transportes y Comunicaciones (MTC), 2024). It is the country's first underground metro system and, as of the end of April 2026, recorded a physical construction progress of 83.12 % (Ministerio de Transportes y Comunicaciones (MTC), 2026). Once completed, Line 2 will connect the district of Ate with Callao, integrating ten districts of Metropolitan Lima and the Constitutional Province of Callao, and will make it possible to complete that journey in approximately 45 minutes, compared with the more than two hours it currently requires by conventional transport, benefiting more than 2.5 million citizens (Ministerio de Transportes y Comunicaciones (MTC), 2026). Likewise, the Ministry of Transport and Communications has advanced the development of Lines 3 (34.8 kilometers) and 4 (23.6 kilometers), both of which obtained technical feasibility in 2020 and are currently under implementation (Ministerio de Economía y Finanzas (MEF), 2025).

More recently, the Urban Transport Authority for Lima and Callao (ATU) presented in 2025 the Urban Mobility Plan (PMU) 2025–2045, a roadmap that, with the support of the Inter-American Development Bank (IDB), guides the transformation of the metropolis's public transport toward an Integrated Transport System (SIT) over the next two decades (Autoridad de Transporte Urbano para Lima y Callao (ATU), 2025). Among its short-, medium-, and long-term projects, the PMU envisages the construction of 90 kilometers of new metro lines and 515 kilometers of segregated corridors, so that the projected network expands from the six

lines contemplated in the original design to nine, additionally incorporating BRT corridors, cable cars, and feeder routes (Autoridad de Transporte Urbano para Lima y Callao (ATU), 2025). It should be noted that the present study focuses on Line 1 of the RBMLC, as it is the network component with the longest operational track record and, therefore, the only one that allows an impact evaluation on urban-mobility outcomes.[1]

From the perspective of the theory of change, the RBMLC articulates a logical chain that links inputs and activities with outputs, intermediate outcomes, and long-term impacts, with the purpose of structurally transforming the urban transport system of Lima and Callao.

At the starting point of this chain are the *inputs*: financial, human, and technical resources allocated to the planning, construction, and maintenance of the rail infrastructure, as well as to the development of urban-traffic management and monitoring systems. These inputs are mobilized through a set of *activities* that encompass pre-investment studies, route design and planning, the construction of the rail infrastructure, the implementation of demand-management systems, the specialized training of operators and drivers, and the communication campaigns aimed at promoting the population's adoption of the system. The intervention also contemplates mechanisms for progressive integration through unified fare systems, technological interoperability, and accessible user-information platforms.

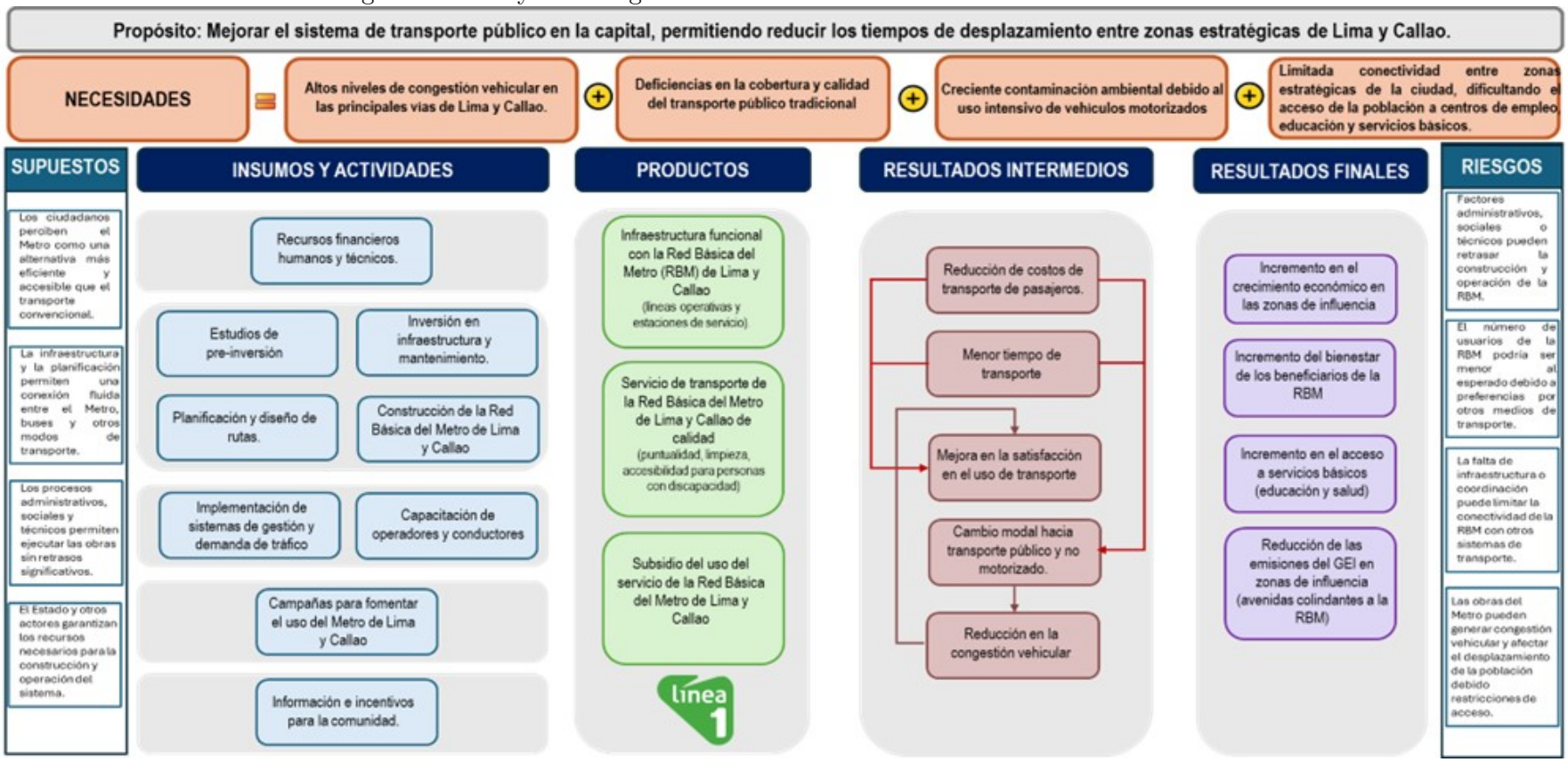
The execution of these activities gives rise to the intervention's central *outputs*: the provision of functional mass transit infrastructure and the delivery of an electric urban transport service with standards of punctuality, frequency, accessibility, and coverage superior to those of the traditional system. Taken together, these outputs aim to consolidate an integrated rail network capable of expanding urban connectivity and of attenuating the structural limitations stemming from the fragmentation and informality of conventional transport.

Building on these outputs, the intervention seeks to generate *intermediate outcomes* linked to the efficiency of urban transport, including the reduction of travel times and transportation costs for users, the alleviation of vehicular congestion, increased satisfaction with the service, and the encouragement of modal shifts toward public and non-motorized transport. These outcomes respond directly to the needs identified in the intervention's diagnosis, in particular the high levels of congestion, the low quality of traditional public transport, and the limited connectivity between strategic areas of Lima and Callao.

Over the long-term horizon, these outcomes are expected to translate into *impacts* of broader scope on urban development and social welfare. Among them are greater efficiency and sustainability of the transport system, expanded population access to basic services, education, and labor markets, the economic dynamization of the system's areas of influence, and the reduction of the environmental externalities associated with conventional motorized transport. Taken together, the RBMLC is conceived as a strategic intervention oriented not only to improving urban mobility, but also to strengthening territorial integration, social inclusion, and environmental sustainability in Metropolitan Lima and Callao. The present study focuses on the intermediate outcomes of this chain, specifically on the effects of Line 1 on users' travel times and transportation costs.

[1]The complete network projected by the PMU 2025–2045 comprises nine metro lines, whose origin and destination alignments are summarized as follows: L1: Villa El Salvador ↔ San Juan de Lurigancho; L2: Ate ↔ Callao; L3: Comas ↔ San Juan de Miraflores; L4: Callao ↔ Ate; L5: Miraflores ↔ San Martín de Porres; L6: Surco ↔ Comas; L7: Ancón ↔ Chorrillos; L8: Barranco ↔ La Molina; L9: Magdalena ↔ San Juan de Miraflores. Source: authors' own elaboration based on Autoridad de Transporte Urbano para Lima y Callao (ATU) (2025).

Figure 3: Theory of Change of the Basic Metro Network of Lima and Callao



Source: Authors' own elaboration.

3 Literature Review

3.1 International Literature

The provision of mass transit systems has the potential to significantly transform the areas located around stations. In Singapore, Diao et al. (2017) found that the Circle Line raised the value of nearby properties by 11%, even before the start of operations, owing to improvements in accessibility and the associated public investment. Similarly, in Medellín, the Metrocable raised land values, although it also generated gentrification processes that displaced low-income residents (Bocarejo et al., 2014). Along these lines, Brinkman and Lin (2026) note that the effects of transport infrastructure reflect a balance between accessibility gains and the local costs associated with its implementation. Although the reduction of travel times and greater access to job opportunities are the main mechanisms of urban benefit, their impacts depend on the territorial context and on the design and location of the infrastructure. In this sense, the international evidence shows that the benefits of mass transit systems may be accompanied by social and territorial challenges.

Beyond these local effects, a broad literature has analyzed the aggregate impacts of transport infrastructure on regional economic development. The reviews by Redding and Turner (2015) and Donaldson (2025) show that these investments generate benefits by reducing commuting costs, expanding market access, and improving productivity. These mechanisms help explain why the expansion of transport networks can generate broader effects on urban and regional growth. In particular, in a meta-analysis of studies on high-speed rail, Cheng and Chen (2021) finds robust evidence of positive impacts on the reduction of travel times, accessibility, and regional economic development. In Germany, Ahlfeldt and Feddersen (2018) shows that HSR systems contributed to the growth of regional GDP, especially at intermediate stations, through agglomeration economies and higher levels of labor productivity. Consistently, the evidence for China shows that the expansion of HSR fostered competitiveness, sectoral specialization, and territorial integration, albeit with heterogeneous effects across cities (Wang et al., 2020; Yang et al., 2022).

Another relevant impact concerns firm dynamics. In Spain, Matas et al. (2020) found that proximity to HSR stations stimulated the creation of new firms, particularly in knowledge-intensive and service sectors. However, these effects were not homogeneous: larger cities were better able to seize the opportunities generated by this infrastructure, whereas smaller ones faced greater difficulties in capitalizing on its benefits.

At the metropolitan scale, the effects are also significant. In Bogotá, Tsivanidis (2026) documented increases in economic activity in the neighborhoods near the stations of the TransMilenio system; the system not only improved travel times but also promoted the growth of previously underserved areas, generating more inclusive improvements in urban welfare. As a complement, Acton et al. (2022) analyzed the effects of BRT systems on property values and found that the design of the infrastructure, for example dedicated lanes versus mixed traffic, significantly influences the observed economic impacts.

Access to efficient transport systems also affects labor-market participation. Kim et al. (2021) evaluated the impact of light rail transit (LRT) systems at 260 stations in 12 metropolitan areas of the United States and found that access to LRT increased labor-force participation by up to 43% relative to areas without access, an effect reflected in a greater number of weeks worked and a higher share of full-time employment, without significant increases in rental costs. Along similar lines, Dong (2017) found improvements in labor participation associated with urban rail systems, albeit with more modest changes in earnings.

From an environmental perspective, transport accounts for close to 25% of global CO_2 emissions (Van den Berg and De Langen, 2017), mainly through the combustion of fossil fuels, and constitutes one of the main sources of greenhouse gases in urban settings. In response, electric rail systems emerge as a cleaner alternative. In their analysis of the Shenzhen metro network, Hu et al. (2024) found that the expansion of the system not only increased the share of residents using the metro, but also reduced car use, contributing to lower urban carbon emissions. This evidence highlights the role of electrified public transport infrastructure in promoting more sustainable urban environments.

Recent evidence reinforces these findings with perspectives specific to lower-middle-income countries, particularly pertinent for the Peruvian case. Using a before-after design applied to the Mumbai Metropolitan Region, Abhijna et al. (2025) documented that 81% of metro users reduced their travel time by an average of 26 minutes per day, allocating the recovered time primarily to family activities, together with significant improvements in perceptions of comfort, reliability, and safety of the service. Along the same lines, Ayaragarnchanakul et al. (2025) analyze the expansion of the Bangkok metro using taxi GPS data and a difference-in-differences framework, and find that the opening of 19 new stations reduced taxi trips by approximately 32.2% in the system's areas of influence, with effects that grow over time and remain robust within catchment radii of up to 800 meters. This last study constitutes a relevant methodological reference for evaluating the effects of the metro on modal substitution and transportation costs in developing cities, dimensions central to the present work. Finally, the international literature warns that sustaining these benefits over time requires comprehensive asset-management frameworks that adequately incorporate future demand, technical performance, and financial planning (Mohammadi et al., 2019), as well as robust strategies for managing service disruptions, including internal contingency protocols and coordination with alternative systems such as replacement buses, in order to preserve operational resilience and user confidence (Zhang et al., 2021). These dimensions, extensively studied in developed economies, remain scarcely explored in the Peruvian context and, more broadly, in developing countries, despite being key elements for advancing toward sustainable, equitable urban transport models adapted to local realities.

3.2 Peruvian Literature

In the Peruvian context, recent studies on the mass transit systems of Metropolitan Lima and Callao, mainly the Metropolitano (BRT) and Line 1 of the RBMLC, highlight relevant impacts on the urban, economic, and environmental development of the city. Among the most important contributions, Scholl et al. (2018) evaluated the impact of the Metropolitano on employment and labor formality. Their results show that the areas near the trunk corridor experienced significant increases in employment rates, hours worked, and labor income, suggesting improvements in job quality. Nonetheless, the areas served only by feeder routes, without direct access to the trunk corridor, obtained far more limited benefits, revealing an unequal distribution of the system's positive effects and the need for policies aimed at reducing the gaps between central and peripheral areas.

As a complement, Oviedo et al. (2019) analyzed how the reduction in travel times associated with the BRT improved access to employment centers for residents of peripheral areas. Although accessibility improved in general terms, the lower-income communities located outside the trunk corridor continued to face significant constraints. This pattern reinforces the idea that the territorial equity of mass transit benefits depends heavily on the location of stations and on network coverage.

This concern with transport affordability has recently been systematized in a comprehensive analysis for Latin America and the Caribbean. Oviedo et al. (2025) examine the theoretical and empirical evolution of the concept of transport affordability in the region and note that disproportionate spending on mobility constitutes a significant source of economic stress and vulnerability for low-income households, with direct implications for social exclusion and access to opportunities. The authors stress that purely monetary conceptualizations of affordability are insufficient to capture the socio-material dimension of (im)mobility, which also depends on trip characteristics, household budgeting practices, and available payment technologies. This perspective is especially pertinent for Lima, where transport-cost constraints overlap with patterns of spatial segregation that limit the access of the most vulnerable households to the benefits of Line 1.

With respect to transport infrastructure and gender, Velásquez (2023) examines how improvements in infrastructure, such as the Metro and the BRT, affect the labor participation of women and men. The study finds that the reduction of commuting times narrows the within-household gender wage gap in dual-earner households, by facilitating women's labor access, and identifies interdependence in the household's mobility decisions, whereby one spouse's transport decisions influence the other's labor decisions. However, these effects are highly localized and vary significantly across municipalities, suggesting that the accessibility and benefits of the system depend on geographic location and household structure. This finding connects directly with the national evidence on the gender gap in transport times, which has widened over the last decade (Céspedes-Reynaga, 2025).

Another dimension addressed is the relationship between transport infrastructure and housing informality. Contreras et al. (2024) analyzed the impact of Line 1 of the Lima Metro on the reduction of housing informality and found that the opening of the first segment in 2011 generated a significant decline in housing informality in the areas near the stations. Nonetheless, these effects were short-lived, with no significant changes in subsequent periods. The identified mechanisms, which include improvements in housing quality and migration, suggest that transport infrastructure affects broader processes of urban transformation, although no clear patterns of migration toward areas adjacent to the metro were detected.

In the educational domain, Alba Vivar (2024) analyzed the impact of transport infrastructure on university enrollment, institutional choice, and early labor-market outcomes. The study shows that the reduction of travel times increased university enrollment, especially at private institutions, and identified gender differences: women tended to enroll in lower-quality private institutions, whereas men more frequently chose public universities. In addition, access to transport infrastructure was associated with higher graduation rates and better labor-market outcomes, particularly among students residing near the stations.

Regarding environmental impacts, Line 1 of the Lima Metro illustrates how electric rail transport can generate positive effects on the environment. Through its electric propulsion system, the Lima Metro emits only 14 grams of CO_2 per passenger-kilometer, considerably below the 68 grams of conventional public transport and the 104 grams of private automobiles. Morales (2024) quantified the environmental benefits of the initiative, highlighting reductions in greenhouse-gas emissions and in air pollutants such as $PM_{2.5}$, NO_x, and SO_2, which confirms the positive environmental effects associated with the start of operations of Line 1.

In sum, the reviewed literature reveals two gaps that motivate the present study. First, although the international evidence on the effects of mass transit is abundant, most studies concentrate on developed economies or on large Asian cities, and rigorous causal evidence for middle-income contexts in Latin America remains limited. Second, the national evidence available for Lima and Callao has focused predominantly on the Metropolitano (BRT) and on

dimensions such as employment (Scholl et al., 2018; Oviedo et al., 2019), gender (Velásquez, 2023), housing (Contreras et al., 2024), or education (Alba Vivar, 2024), whereas the effects of Line 1 of the Metro on the core mobility outcomes, that is, users' travel times and transportation costs, have not been quantified causally. The present study contributes to closing both gaps by estimating, through an instrumental-variables strategy that addresses non-random selection into system use, the impact of Line 1 on the travel times and costs of its users. In this way, the work contributes to the urban-economics and infrastructure literature empirical evidence on the returns to mass transit systems in cities of emerging economies, and provides relevant inputs for the design and expansion of sustainable urban-mobility policies in Peru.

4 Empirical Strategy

4.1 The Endogeneity Problem in Mode Choice

This study seeks to estimate the causal effect of using Line 1 of the RBMLC on trip characteristics: in-vehicle time, monetary cost, effective speed, and waiting time. An Ordinary Least Squares (OLS) estimate of this effect would, however, be inconsistent, because the decision to use the metro is not random. Users self-select according to observable and unobservable attributes, such as a preference for speed or an aversion to traffic, which also determine trip outcomes (Wooldridge, 2010). The treatment is thus correlated with the error term, so that the estimator would capture, in addition to the causal effect, the selection bias.

The magnitude of this bias is not a purely theoretical concern. Guerrero et al. (2021), using a control-function approach applied to mode-choice models in Valparaíso (Chile), document that ignoring endogeneity overestimates the subjective value of travel time by between 26% and 49% and underestimates modal elasticities by between 33% and 75%. That is, failing to correct for self-selection leads to erroneous conclusions about the social benefits of transport infrastructure.

In Metropolitan Lima the problem is aggravated for two reasons. First, the marked socioeconomic heterogeneity of users and the spatial distribution of stations cause those who use the metro to differ systematically from those who do not. Second, a mechanism of residential self-selection operates: some households choose their location taking into account access to the system, so that the correlation between proximity to stations and public-transport use may reflect pre-existing preferences rather than the causal effect of the infrastructure, as Tan et al. (2025) show for metro expansions in China.

Faced with this type of endogeneity, the literature on transport-infrastructure evaluation recommends explicit causal-identification strategies. Graham (2025) reviews the available methods, such as instrumental variables, difference-in-differences, and regression discontinuity, and concludes that, when the treatment is endogenous and experimental variation is unavailable, instrumental variables constitute an appropriate strategy for recovering consistent causal estimates.

Following this literature, the present study adopts an instrumental-variables approach with high-dimensional fixed effects (IV-HDFE). The instrument, the spatial proximity of the trip to the network, isolates the exogenous variation in Line 1 use from the unobservable factors that simultaneously affect mode choice and trip outcomes. Under the assumptions of relevance, exogeneity, and monotonicity, the estimator identifies a local average treatment effect (LATE) on the *compliers*: travelers whose decision to use the metro responds to their proximity to the network (Imbens and Angrist, 1994). The estimated effects are therefore informative about this subpopulation, precisely the one relevant for gauging the benefits of expanding the

system's coverage.

4.2 Identification Strategy: Spatial Instrumental Variables

To isolate the exogenous variation in mode choice, this study implements an Instrumental Variables (IV) strategy that exploits the quasi-experimental variation generated by the spatial configuration of the metro infrastructure relative to the origin and destination points of users' trips within the city. Specifically, we construct the binary instrument Z_{iodt}, which takes the value of 1 if the sum of the Euclidean distance from the trip's origin to the nearest station, and from the alighting station to the final destination, is less than or equal to 1 km; and 0 otherwise. The identification logic of this instrument rests on the theory of access costs and last-mile friction, extensively documented in urban-transport economics. Classical mode-choice models establish that the probability of adopting a transport mode is a decreasing function of the generalized costs of access, among which the time and effort of walking to the boarding point are determinant components (Wooldridge, 2010).

Within this framework, the physical configuration of the rail network—defined exogenously by engineering and public-investment decisions made decades before the analysis period—generates a spatial discontinuity in the costs of access to the system: users whose origins and destinations fall within the operational pedestrian radius ($\leq$ 1 km) face substantially lower friction costs than those located beyond that threshold. This discontinuity, being the product of technical decisions rather than of users' mobility preferences, constitutes a source of quasi-experimental variation appropriate for causal identification.

The 1 km threshold employed is not arbitrary. The international literature on pedestrian accessibility to mass transit systems converges on the finding that the maximum distance an average user is willing to walk to access a station ranges between 400 meters and 1.2 kilometers, depending on the urban context (Daniels and Mulley, 2013). In Latin American megacities with high residential density and fragmented road networks—conditions characteristic of Metropolitan Lima—studies such as Oviedo et al. (2019) document pedestrian-catchment patterns consistent with this threshold. The choice of 1 km *cumulative* (the sum of the distance to the origin plus the distance to the destination) is moreover conservative: it ensures that users classified as "proximate" can complete both ends of their trip on foot, without the need for additional approach modes that would introduce unobserved transaction costs.

Relevance Assumption (First Stage)

The relevance condition requires that the instrument predict the endogenous variable—Line 1 use—sufficiently strongly once the observable covariates and fixed effects are controlled for. The theoretical mechanism is direct: geographic proximity reduces friction and last-mile costs, monotonically increasing the expected net utility of the metro relative to its substitutes. As a consequence, users whose combined origin and destination lie within the operational pedestrian radius have a significantly higher probability of choosing Line 1 over conventional transport alternatives. The first-stage estimates empirically confirm this prediction. For the baseline specification pooling the entire population sample, the Kleibergen-Paap Wald $rk\ F$ statistic reaches a value of 29.77, comfortably exceeding both the classical rule of thumb of 10 established by Stock and Yogo (2005) and the stricter thresholds proposed by the recent literature (Lee et al., 2022). In addition, the robustness of the inference is confirmed by the Anderson-Rubin Wald test, which consistently rejects the null hypothesis in the main specifications ($p < 0.01$), guaranteeing the validity of the second-stage coefficients regardless of the instrument's strength in small samples.

Exclusion Restriction: Arguments and Potential Threats

The exclusion restriction requires that spatial proximity to the Line 1 stations affect trip outcomes solely and exclusively through its effect on the probability of using the metro. This assumption, although not directly testable, can be assessed through theoretical arguments and indirect evidence. First, the location of the stations responds to a logic of rail engineering and high-demand corridor planning defined between the 1970s and 1990s, well before the analysis period (Corporación Andina de Fomento (CAF), 2015; ATU, 2023). This historical predetermination implies that the spatial distribution of the stations is, to a large extent, independent of the travel preferences and mobility conditions observed in 2023; consequently, the variable Z_{iodt} captures variation in physical accessibility to the system that is not contaminated by contemporaneous demand factors.

Nonetheless, three potential threats deserve attention. The first is *residential sorting*: if households with systematically different preferences for speed or savings actively select to reside near the stations, spatial proximity would be correlated with unobservable characteristics that independently affect trip outcomes. This concern is empirically relevant—Tan et al. (2025) document, for metro expansions in China, that such a relationship may partly reflect households' pre-existing preferences—but the IV-HDFE model mitigates it directly through origin-district (α_o) and destination-district (α_d) fixed effects, which absorb the unobservable spatial heterogeneity—road-infrastructure endowment, local topography, and the socioeconomic composition of neighborhoods—isolating the variation in proximity within each origin-destination pair.

The second threat arises if the presence of stations per se improves the conditions of the surrounding road environment—investments in sidewalks, signage, or traffic reorganization—affecting travel times even for those who do not use the metro. The available evidence suggests that these spillover effects are limited in the Line 1 corridor: Contreras et al. (2024) document that the impacts on the urban environment are localized and short-lived, without systematic road transformations independent of the transport mode. In addition, the departure-hour fixed effects (τ_t) control for temporal congestion shocks that could differentially affect users near and far from the stations.

The third threat is that spatial proximity may be correlated with structurally distinct trip types—shorter trips or trips in less congested areas—that generate differences in outcomes independently of the mode used. This possibility is addressed by including in the vector X_i controls for the total trip distance, the number of transfers, and the departure hour, which capture the main sources of heterogeneity in trip structure.

From a design standpoint, the strategy can be interpreted as a fuzzy spatial discontinuity in the probability of metro use at the cumulative-access threshold of 1 km. Under this reading, the key assumption is the comparability between trips immediately below and above the threshold, in line with the literature on geographic discontinuities (Keele and Titiunik, 2015; Dell, 2010). Likewise, the validity of the conclusions does not rest on strictly perfect exogeneity: the plausibly-exogenous instruments framework of Conley et al. (2012) allows the causal effect to be bounded under moderate violations of the exclusion restriction, while the zero-first-stage test of van Kippersluis and Rietveld (2018) offers a way to test empirically for the absence of a direct effect of the instrument in subpopulations where the metro is not a feasible substitute.[2] The stability of the coefficients under the progressive inclusion of fixed effects and controls

[2]The formal implementation of these sensitivity tests—bounds of Conley et al. (2012), estimation under sample trimming (*trimming*) to windows around the 1 km threshold, and the zero-first-stage test—constitutes the natural extension of the design and is incorporated as an identification-robustness agenda.

additionally constitutes indirect evidence that selection on unobservables does not dominate the results (Oster, 2019).

Comparison with Alternative Instruments in the Literature

The choice of spatial proximity to stations as an instrument for mass transit use is embedded in a consolidated methodological tradition in transport economics, though with important variants depending on the context of application.

The strategy closest to the design of this study is that employed by Tsivanidis (2026) to evaluate the impact of the TransMilenio system in Bogotá, who uses distance to BRT stations as a source of exogenous variation in access to the system. Analogously, Ayaragarnchanakul et al. (2025), in their analysis of the expansion of the Bangkok metro using taxi GPS data, exploit the opening of new stations as an exogenous spatial *shock* in a difference-in-differences design. Both studies share with the present work the intuition that the physical configuration of the infrastructure generates discontinuities in access that are independent of individual mobility preferences.

An instrumental alternative frequently employed in the literature is the use of historical plans or unimplemented transport-infrastructure projects as an instrument for the current network. Duranton and Turner (2012) pioneered this logic in the context of interurban highways in the United States, using the 1947 highway plan and 1898 railroad maps as instruments for the 1983 road network. Following a similar strategy, Gonzalez-Navarro and Turner (2018) employ the historical predetermination of rail networks to identify the effects of metro systems on urban growth in a global sample of 632 cities. This strategy, while successfully exploited in contexts with high availability of historical data, presents operational limitations in the case of Lima: the original blueprints of the RBMLC have undergone multiple revisions since 1973, and the availability of high-quality georeferenced data for the discarded historical alignments is limited.

Another alternative is the use of topography—such as slope or the presence of natural barriers—as an instrument for the road or transport network, also following the logic of Duranton and Turner (2012). However, since Line 1 operates for the most part at grade and on an elevated viaduct along a relatively flat north-south corridor, topographic variation does not constitute a sufficiently powerful source of identification in this specific context.

Relative to these alternatives, the cumulative-proximity instrument presents clear comparative advantages: it is directly computable from high-precision georeferenced data, its transmission mechanism is theoretically sound and empirically verifiable in the first stage, and its cumulative design—summing the distances at both ends of the trip—captures the total access friction to the system more completely than instruments based solely on distance to the origin.

4.3 The IV-HDFE Model Specification

To mitigate possible violations of the exclusion restriction stemming from *residential sorting* (the tendency of households with particular characteristics to locate near the stations) and to isolate spatial and temporal shocks, the IV model is enriched by the inclusion of High-Dimensional Fixed Effects (HDFE).

The system of simultaneous equations is defined as follows. The first-stage equation (mode selection) is:

$$Metro_{iodt} = \pi_0 + \pi_1 Z_{iodt} + X_i'\gamma_1 + \alpha_o + \alpha_d + \tau_t + \nu_{iodt} \tag{1}$$

The structural second-stage equation (trip outcomes) is:

$$Y_{iodt} = \beta_0 + \beta_{LATE}\widehat{Metro}_{iodt} + X_i'\gamma_2 + \alpha_o + \alpha_d + \tau_t + \epsilon_{iodt} \tag{2}$$

where:

- Y_{iodt} denotes the outcome variable of interest (in-vehicle time, total cost, speed, or waiting time) for individual i, traveling from origin district o to destination district d, with departure hour t.
- $Metro_{iodt}$ is the binary endogenous variable that takes the value of 1 if the individual used Line 1 of the RBMLC in some stage of the trip.
- $\widehat{Metro}_{iodt}$ is the predicted probability of using the train obtained from the first stage.
- Z_{iodt} is the spatial-proximity instrument (≤ 1 km).
- X_i is a vector of individual- and trip-level control covariates that includes: total trip distance, age, age squared, sex, socioeconomic status (NSE), student status, and the number of public-transport transfers made.
- α_o and α_d are origin- and destination-district fixed effects, respectively, which absorb the unobservable spatial heterogeneity, such as road-infrastructure endowment or local topography.
- τ_t captures departure-hour fixed effects, controlling for the structural variation of congestion and urban-mobility demand over the course of the day.
- ϵ_{iodt} and ν_{iodt} are the stochastic error terms.

Given the nature of the Urban Mobility Plan (PMU 2023) data, the transport decisions of individuals living in the same dwelling may be correlated. Therefore, the stochastic standard errors (ϵ_{iodt}) are clustered at the household level to ensure valid statistical inference. Estimation is carried out on a restricted subsample of trips that use transport modes potentially substitutable for Line 1, excluding purely non-motorized trips or trips made in exclusive private vehicles.

Separately, as a robustness exercise, we assess whether the results depend on the administrative definition of the spatial units of origin and destination. To this end, we re-estimate the model using homogeneous spatial units constructed through an H3 hexagonal grid, instead of administrative districts.

This specification allows us to assess whether the estimate of the effect of Line 1 use is sensitive to the spatial aggregation of trips. In particular, we consider interacted fixed effects between the origin and destination H3 units, keeping the hourly fixed effects and the set of sociodemographic controls of the main specification.

In this spatial-robustness exercise, the estimated effect on speed is 12.57 km/h under the H3-R7 grid, compared with 13.58 km/h in the corresponding district specification. Both magnitudes are close to each other and consistent in sign and significance with the 11.83 km/h

effect reported in the main aggregate specification (Table 5), which indicates that the result is not particularly sensitive to the spatial definition used.

The H3-R7 specification implies a reduction in the sample owing to the structure of the fixed effects and the elimination of singleton observations. Its results are therefore interpreted as complementary robustness evidence and not as a substitute for the main specification.

4.4 Causal Interpretation (LATE) and General-Equilibrium Effects

The estimated parameter of interest, β_{LATE}, does not represent the population Average Treatment Effect (ATE), but the Local Average Treatment Effect (LATE). This implies that the evaluation results quantify the social return (savings in time and cost) of Line 1 strictly for the group of *compliers*: those users who decided to use the metro induced by their spatial proximity (≤ 1 km), and who would have chosen an alternative mode in the absence of such proximity.

In addition, because of the use of a post-intervention cross-sectional design, it is essential to recognize that the model operates under general-equilibrium dynamics. The entry into operation of Line 1 generated a massive supply *shock* that endogenously reorganized the traditional public-transport market along the axes of influence (changes in frequencies, routes, and fares of combis, coasters, and corridors). Therefore, the counterfactual implicit in this estimation captures the relative benefit of Line 1 vis-à-vis the substitute alternatives in their current equilibrium state, absorbing the *spillovers* that the metro itself generated on the rest of the urban-mobility ecosystem.

4.5 Heterogeneity Analysis and Spatial Sensitivity

The main estimate is disaggregated to capture the heterogeneity of the effects along two dimensions. The first is temporal: we distinguish between the morning and evening peaks, the off-peak periods, and the nighttime schedule, which allows us to assess whether the metro's benefits vary with the congestion conditions of the road network. The second is socioeconomic: we stratify by household socioeconomic status (High, Middle, and Low), given that the opportunity cost of time and budget constraints differ substantially across income groups.

As a sensitivity exercise, we relax the spatial threshold that defines the proximity instrument. The baseline specification considers a cumulative radius of 1 km between the origin, the network, and the destination; for the middle-income stratum we additionally evaluate an expanded radius of 2 km. This variation is not arbitrary: it responds to the fact that the metro's functional area of influence depends on the "last-mile" approach modes, whose reach varies with income level. Middle-income households tend to combine Line 1 with short-distance, low-cost access modes, which extends their effective catchment radius beyond direct walking. Contrasting both thresholds allows us, on the one hand, to verify the robustness of the results to the definition of the instrument and, on the other, to assess whether its predictive power improves when the radius is adjusted to the access patterns observed in this subsample.

The selection of the instrument's radius for each stratum is guided by the strength of the first stage and not by the second-stage result; Table 4 transparently reports both thresholds (1 km and 2 km) for the three strata, so that the reader can assess the sensitivity of the effect to this choice. The focus of the analysis is the qualitative pattern of progressivity and not the magnitude of each coefficient.

5 Data and Variables

5.1 Data Source

The empirical analysis of this study is based on the microdata of the Household Mobility Survey of the Urban Mobility Plan for Lima and Callao (PMU) 2023, prepared by the Urban Transport Authority for Lima and Callao (ATU) with the support of the Inter-American Development Bank (IDB) (ATU, 2023). This survey constitutes the largest-scale mobility study for Lima and Callao since the JICA origin-destination survey in 2012 (JICA, 2012), and it is part of the fieldwork that underpinned the formulation of the PMU 2025–2045 (Autoridad de Transporte Urbano para Lima y Callao (ATU), 2025).

Data collection was carried out between September and November 2023 and targeted residents of private households in Metropolitan Lima and the Constitutional Province of Callao. The sample design covers 50 districts of both jurisdictions and records the trips of persons aged 5 and over. The survey captures the trips made on typical weekdays, Monday through Friday, excluding Saturdays, Sundays, and holidays, which focuses the analysis on the patterns of regular mobility associated with work, study, and other everyday activities. According to the PMU diagnosis, around 24 million trips are made daily in Lima and Callao, with peak demand concentrated in the morning band from 7:00 to 9:00 a.m. and work as the main trip purpose, followed by study (ATU, 2023).

A central feature of the survey is its hierarchical structure at three nested levels, which makes it possible to link each trip with the characteristics of the person who makes it and of the household to which they belong:

- **Household level**: records the attributes of the dwelling and the household, including its geographic location, socioeconomic status (NSE), monthly income, housing-tenure regime and, relevant for this study, the household's vehicle ownership and fleet (number, age, and type of vehicles, from bicycles and motorcycle taxis to private automobiles).
- **Individual level**: contains the sociodemographic characteristics of each household member, such as age, sex, kinship, educational level, occupation, disability status, driver's-license holding, and work arrangement (in-person, hybrid, or remote).
- **Trip level**: constitutes the unit of analysis of the study. Each record documents a trip reported by respondents during the day prior to the interview, with information on the (georeferenced) origin and destination, the departure and arrival times, the trip purpose, and a detailed description of each of the stages of the journey. For each stage, the transport mode used (up to eight chained modes per trip), the travel time, the waiting time, and the monetary cost are recorded, which makes it possible to accurately reconstruct the total times and costs of the transport chain.

Since each level incorporates its own expansion factor, the results are representative of the entire population of Metropolitan Lima and Callao. In particular, the trip-level expansion factors allow the sample to be projected to approximately 20 million trips per day in the metropolitan area, a figure consistent with the official estimates of the PMU diagnosis (Autoridad de Transporte Urbano para Lima y Callao (ATU), 2025).[3]

[3]The survey classifies each trip according to its time of day into morning peak (HPM), off-peak (valle), evening peak (HPT), and night, which enables control for congestion patterns over the course of the day.

5.2 Construction of the Analytical Sample

The identification strategy of this study compares the mobility outcomes of Line 1 users with those of a comparison group of travelers who do not use it. The validity of this comparison rests, critically, on the counterfactual being composed of trips genuinely substitutable by the metro. A trip made entirely on foot over three blocks, or one made by private car across the city, does not constitute a reasonable alternative to using Line 1 and, if included, would introduce composition bias by mixing structurally distinct mobility decisions. For this reason, the analytical sample is restricted to the universe of trips *potentially substitutable* by the metro, through the sequential application of the following criteria.

(i) Exclusion of non-comparable modes. We discarded trips made entirely by non-motorized modes (walking and cycling), those made exclusively by private transport (private automobile and motorcycle), and the atypical or marginal modes within the sample (trucks, trailers, and scooters). The comparison group is thus composed of users of mass and conventional public transport (SIT corridors, Metropolitano, buses, minibuses, combis, and shared taxis), which represent the effective alternative to the electric train.

(ii) Operational trip restrictions. We removed trips of very short distance (less than 1 km) or with in-vehicle time under 10 minutes. The logic is that these trips do not justify, in terms of access time and cost, the use of a metro system, so their inclusion would distort the comparison of times and costs across modes.

(iii) Consistency and timing filters. We retained only trips with a number of transfers between 1 and 4, with a strictly positive monetary cost, and whose start was recorded within the regular operating hours of the city's transport system (between 05:00 and 22:59). These criteria clean out records with inconsistent or extreme values (for example, trips with zero cost or started at dawn, outside the metro's service hours) that would compromise the comparability of the control group.

After the sequential application of these filters, the final sample consists of 12,684 motorized urban trips. All estimations employ the analytical expansion factors provided by the PMU (ATU, 2023), so that the results retain their representativeness with respect to the population of motorized urban trips of Lima and Callao. It should be stressed that these filters redefine the inference population: the estimated effects are valid for the universe of trips potentially substitutable by the metro, and should not be mechanically extrapolated to the set of all trips in the metropolis.

Table 1: Description of Variables Used in the Estimations

Classification	Variable / Symbol	Description
Outcome (Dependent) Variables		
	In-vehicle time (*tbordo*)	Net time (in minutes) that the user spends inside the transport vehicle or vehicles, excluding walking and waiting times.
	Trip cost (*costo_viaje*)	Total monetary outlay (in Soles) reported by the user to complete the entire journey.
	Transport speed (*vel_transporte*)	Effective speed (in km/h) computed as the ratio of the Euclidean trip distance to in-vehicle time.
	Waiting time (*tespera*)	Total time (in minutes) spent on platforms or at stops before boarding the transport unit.
Endogenous Variable (Treatment)		
	Metro use (*modalidad_tren*)	Binary variable that takes the value of 1 if the respondent reported using the "electric train" (Line 1) in any stage of the trip (from the first to the eighth transport mode), and 0 if the trip was completed using exclusively other mass or conventional public-transport modes (SIT, combis, coasters, and shared taxis).
Instrumental Variable		
	Cumulative proximity (*rank1k_dtotal*)	Binary spatial variable that takes the value of 1 if the sum of the straight-line distance from the trip origin to the nearest boarding station, plus the distance from the alighting station to the final destination, is less than or equal to 1 km. For the spatial-sensitivity analysis by socioeconomic status, the expanded threshold of 2 km (*rank2k_dtotal*) is additionally used.
Control Variables (X_i)		
	Trip distance (*distancia_viaje*)	Direct geodesic distance between origin and destination (in km).
	Vehicular transfers (*trasbordos_publico*)	Number of vehicular transfers made (entered as a categorical variable).
	Passenger age (*edad*, $edad^2$)	Passenger age, including its quadratic term to capture non-linear effects.
	Respondent sex	Binary variable identifying female passengers.
	Socioeconomic stratum (NSE)	Categorical recoding of the household's socioeconomic status (High, Middle, and Low) to control for budget constraints and the opportunity cost of time in mode choice.
Fixed-Effects Variables (HDFE)		
	Origin district (α_o)	Categorical identifier (51 categories) to absorb the unobservable geographic heterogeneity and frictions at the point of departure.
	Destination district (α_d)	Categorical identifier (51 categories) to absorb the unobservable geographic heterogeneity and frictions at the point of arrival.
	Departure hour (τ_t / *hh_salida*)	Categorical identifier (18 time bands) to control for aggregate vehicular-congestion shocks that affect the road network of Metropolitan Lima transversally.

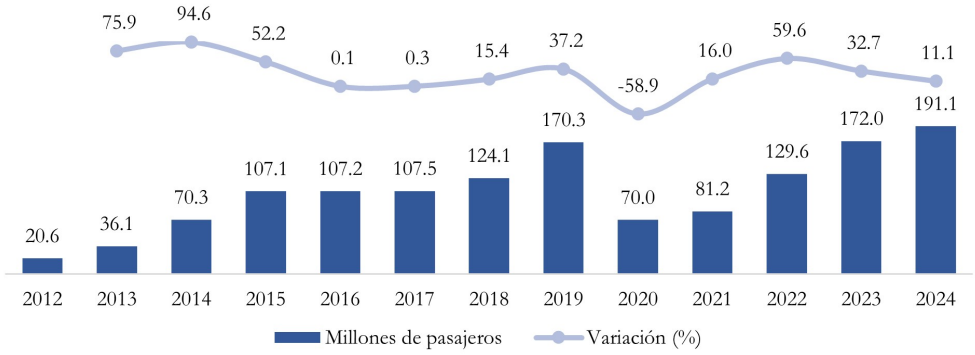
Source: Authors' own elaboration based on the PMU 2023 Survey (ATU, 2023).

6 Descriptive Statistics

Line 1 of the Basic Metro Network of Lima and Callao (RBMLC) has experienced sustained growth in the number of passengers carried since the start of its operations in 2012. Between 2012 and 2019, the number of users increased from 20.6 million to 170.3 million passengers, respectively. This growth was associated with the expansions in infrastructure and equipment aimed at strengthening transport capacity and improving service quality, such as the opening of the second segment in 2014, the expansion of the train fleet, and the incorporation of new units and cars in 2018 and 2019.

In this context, following the significant reduction in the number of passengers observed during the health and economic crisis of 2020 and 2021, the flow of Line 1 users exceeded pre-pandemic levels toward the end of 2023 (172 million passengers) and 2024 (191 million passengers).

Figure 4: Evolution of passenger traffic, 2012–2024

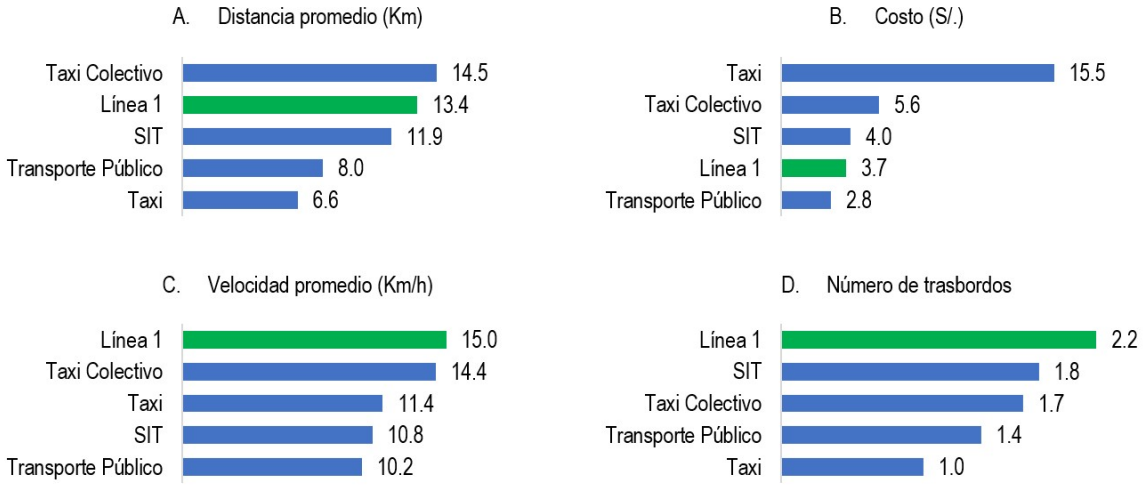


Note: Millions of passengers and percentages.
Source: Urban Transport Authority for Lima and Callao (ATU, 2024).
Authors' own elaboration.

According to the PMU Urban Mobility Survey, Line 1 of the RBMLC consolidates itself as a relatively efficient public-transport mode, characterized by a higher average speed and an affordable trip cost relative to other urban-transport alternatives. In 2023, trips that included Line 1 reached an average speed of 15.0 km/h, higher than that recorded for modes such as shared taxi (14.4 km/h), taxi (11.4 km/h), and the Integrated Transport System (10.8 km/h), composed of the Metropolitano and the road corridors.

Likewise, although trips made using Line 1 involved a greater average distance traveled (13.4 km) and a larger number of transfers (2.2 on average), the total trip cost (S/ 3.7) was considerably lower compared with modes of similar speed, such as taxi (S/ 15.5) and shared taxi (S/ 5.6).

Figure 5: Cost, distance, speed, and number of transfers of trips in Metropolitan Lima, by transport mode share, 2023

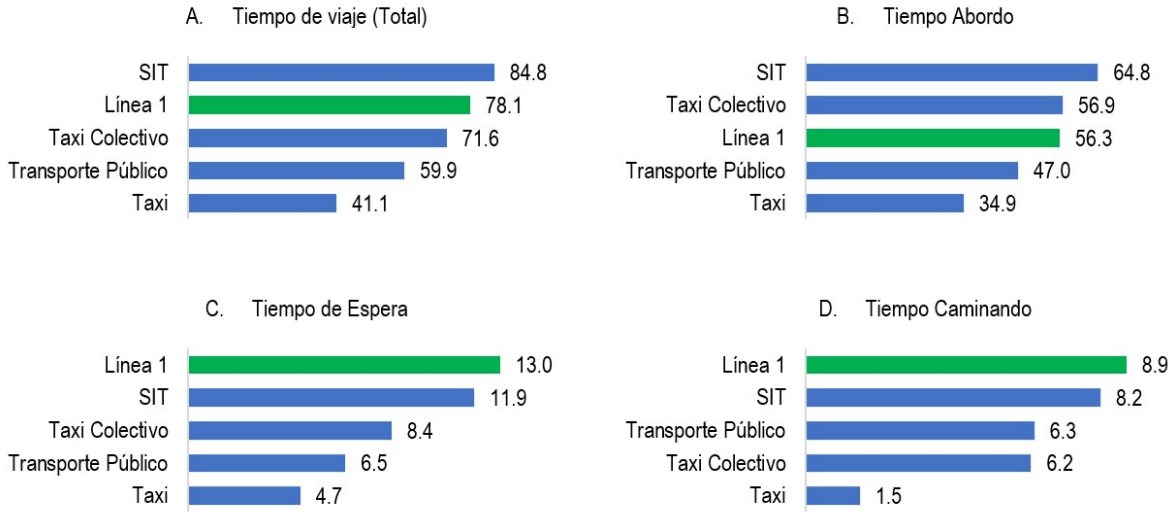


Note: The Integrated Transport System (SIT) includes the road corridors and the Metropolitano. Speed considers the distance between the trip's origin and final destination, as well as the in-vehicle time of the transport mode, excluding waiting and walking time.
Source: PMU Urban Mobility Survey – ATU (2023).
Authors' own elaboration.

Separately, the PMU Urban Mobility Survey shows that, in 2023, Line 1 users recorded higher average waiting times (13.0 minutes) and walking times (8.9 minutes) compared with those who used other transport modes. In particular, 28.0% of the total travel time on Line 1 corresponded to waiting or walking time, a share higher than that observed for the Integrated Transport System (23.7%), the shared taxi (20.4%), and conventional public transport (21.4%).

These results suggest that, despite being a relatively efficient and lower-cost transport alternative, Line 1 faces challenges associated with the complementary times of the trip, particularly in the access and waiting stages. This could reflect limitations linked to demand management and to the high volume of passengers the system serves.

Figure 6: Trip times in Metropolitan Lima, by transport mode share, 2023



Note: Values expressed in minutes. The Integrated Transport System (SIT) includes the road corridors and the Metropolitano.
Source: ATU (2023).
Authors' own elaboration.

7 Empirical Results

7.1 First Stage and Validation of the Spatial Instrument

Before analyzing the social returns to Line 1, it is imperative to validate empirically the soundness of the identification strategy. Table X (see Appendix) reports the first-stage estimates of the IV-HDFE model, where we assess the ability of cumulative spatial proximity ($rank1k_dtotal$) to predict train use.

The results confirm that living and traveling to a destination located 1 km or less from the Line 1 stations positively and highly significantly increases the probability of using the metro over other public-transport alternatives. For the baseline specification pooling the entire population sample, the Kleibergen-Paap Wald rk F statistic reaches a value of 29.77. By comfortably exceeding the classical rule of thumb of 10 established by Stock and Yogo (2005) and the stricter recent inference thresholds (Lee et al., 2022), the presence of a weak-instruments problem is formally ruled out.

In addition, the robustness of the inference is confirmed by the Anderson-Rubin Wald test, which robustly rejects the null hypothesis, guaranteeing the validity and significance of the second-stage coefficients regardless of the instrument's strength. Regarding the validity of the instruments, since the empirical specification employs a single excluded instrument ($rank1k_dtotal$) to instrument a single endogenous variable ($modalidad_tren$), the model is exactly identified. Consequently, the Hansen J statistic takes a value equal to 0.000, which is mathematically expected in just-identified settings and confirms that the instrument provides a unique and sufficient source of variation for the system of equations.

Taken together, this set of results validates that the spatial quasi-experimental design provides sufficient exogenous variation to estimate the Local Average Treatment Effect (LATE) in the second stage with consistent standard errors.

7.2 Aggregate Effects on Trip Efficiency (LATE)

Table 2 presents the second-stage estimates of the IV-HDFE model for the four aggregate outcome variables: in-vehicle time, monetary cost, speed, and waiting time. Having satisfactorily passed all the identification and instrumental-robustness tests (Kleibergen-Paap, Anderson-Rubin, and Hansen J), these results carry causal validity. The estimates reveal that the operation of Line 1 of the RBMLC confers substantial structural benefits on *complier* users, although it introduces a logistical *trade-off* characteristic of high-demand underground or segregated mass transit systems.

First, the results provide evidence of a large time saving that is statistically significant at the 1% level. Users who choose Line 1 induced by their proximity experience an average reduction of 27.80 minutes in the net in-vehicle time relative to the counterfactual of substitute transport. Correspondingly, the effective transport speed increases by 11.83 km/h, evidencing the operational advantage of an exclusive right-of-way over mixed road networks exposed to chronic congestion.

On the economic dimension, the heavily subsidized flat fare of S/ 1.50 translates into financial relief for users, since Line 1 use generates an average saving of S/ 1.23 per trip, statistically significant at the 1% level. This result reveals that the economic benefit is not homogeneous: it is pronouncedly concentrated among users of lower socioeconomic status (S/ 2.03) and among trips during the evening peak (S/ 2.39), positioning Line 1 not only as a mobility solution but as an implicit instrument of income redistribution.

Nonetheless, these positive returns in time and money are accompanied by a non-pecuniary cost in the form of access frictions. The model estimates a significant increase of 2.71 minutes ($p < 0.10$) in pre-boarding waiting times. This coefficient, although marginally significant in the aggregate sample, captures the diseconomies of scale at highly saturated stations (access control, card top-ups, platform walking times, and fixed service intervals), configuring a clear *trade-off*: users tolerate a penalty in waiting and access time in exchange for significantly faster, cheaper, and more predictable in-vehicle trips.

7.3 Time-of-Day Dynamics and General-Equilibrium Effects on Congestion

The aggregate average effect, though positive, conceals substantial intra-day heterogeneity. The urban-transport ecosystem in Lima and Callao is characterized by severe fluctuations in road congestion and in the supply of traditional transport services over the course of the day. Estimating the model segmented by time bands—whose specifications also consistently satisfy the robustness and instrumental-identification tests (Kleibergen-Paap, Anderson-Rubin, and Hansen J)—reveals how the benefits of Line 1 vary as a function of the equilibrium state of the substitute market.

The Metro as a shield against congestion (Peak Hours)
During the daytime demand peaks, the train's segregated right-of-way maximizes its relative

Table 2: Aggregate Effects of Line 1 Use on Trip Efficiency

Variables	(1) In-vehicle Time (min)	(2) Trip Cost (S/)	(3) Speed (km/h)	(4) Waiting Time (min)
Line 1 Use	**-27.80*** **	**-1.23*** **	**11.83*** **	**2.71***
	(4.59)	(0.45)	(3.30)	(1.45)
Diagnostic Statistics				
Observations (N)	12,684	12,684	12,684	12,684
K-P Wald rk F-stat	29.77	29.77	29.77	29.77
Anderson-Rubin Wald test (p-value)	0.000	0.007	0.013	0.046
Hansen J statistic (p-value)	Just-identified	Just-identified	Just-identified	Just-identified
FE: Origin, Destination, Hour	Yes	Yes	Yes	Yes

Note: Household-clustered standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. "Just-identified" indicates that the model is exactly identified ($L = K$), so the Hansen overidentification test does not apply. Excluded instrument: cumulative proximity ≤ 1 km.

value. In the Morning Peak (HPM, 7:00-8:59) and the Evening Peak (HPT, 17:00-19:59), metro use guarantees reductions in in-vehicle time of 22.7 and 22.6 minutes, respectively. In monetary terms, the HPT reports the largest financial relief of the day, with a saving of S/ 2.39 per trip. This peak in savings suggests that, in hours of maximum saturation, the transport alternatives (such as taxis or shared taxis) apply dynamic fares or congestion surcharges from which the metro's flat fare protects the user.

However, it is precisely in these time bands that the logistical *trade-off* becomes most severe. The waiting time to board the train experiences its greatest penalties, reaching an increase of 10.9 minutes during the HPT relative to conventional transport modes. The saturation of platforms and the exceedance of train capacity impose rationing through physical queues, partially absorbing the benefit of the time saved in-vehicle.

Off-peak Hours and the Nighttime Anomaly

In the daytime off-peak periods (9:00-16:59), the train's speed advantage stabilizes (+14.6 km/h) and the waiting frictions moderate. Nonetheless, the most counterintuitive and revealing results emerge in the Nighttime Off-peak (20:00-22:00).

During the nighttime band, Line 1 use reduces waiting time by 6.6 minutes (a complete reversal relative to the rest of the day) and reaches its maximum in-vehicle time saving (35.0 minutes). This responds to a general-equilibrium effect of supply: at night, the frequency of combis and traditional buses drops drastically, increasing the waiting and travel times of the substitute modes. Faced with this contraction of the traditional private supply, the regulated and continuous operation of Line 1 becomes extremely efficient.

In this same nighttime band, the model reports a positive and significant monetary effect: using the metro is on average S/ 0.78 more expensive than the available counterfactual. This result is consistent with the economics of Lima's urban transport. First, the dissipation of congestion reduces the surcharges applied by informal modes and taxis during peak hours, while the metro's flat fare remains constant, which explains the S/ 2.39 saving estimated in the HPT. Second, the fractioned structure of informal transport, characterized by per-passenger competition and short, negotiable per-segment fares, can make a single informal trip cheaper than the fixed fare of the mass transit system (Bielich Salazar, 2009); this pattern of low-cost, flexible-supply services is characteristic of informal paratransit in developing cities (Cervero

and Golub, 2007).

Added to this is the composition of nighttime trips, which may be shorter and more discretionary, making a flat fare relatively costly relative to a distance-related informal fare. Finally, the result may reflect a money-time trade-off: although the nighttime user pays an additional S/ 0.78, they save on average 35 minutes in-vehicle, which suggests a willingness to pay for speed, predictability, and safety relative to more uncertain informal alternatives during the night. Taken together, these mechanisms suggest that Line 1 loses relative price competitiveness precisely when it has the greatest idle capacity, which provides an additional justification for the differentiated-fare policy proposed in Section 9.

Table 3: Heterogeneous Effects by Time of Day

	(1) Morning Peak	**(2) Daytime Off-peak**	**(3) Evening Peak**	**(4) Nighttime Off-peak**
Panel A: In-vehicle Time (minutes)				
Line 1 Use	-22.72***	-25.48***	-22.64*	-34.99***
	(5.88)	(6.80)	(11.97)	(6.00)
A-R test (p-value)	0.005	0.018	0.060	0.001
Panel B: Trip Cost (S/)				
Line 1 Use	-0.49*	-0.05	-2.39*	0.78**
	(0.26)	(0.44)	(1.31)	(0.36)
A-R test (p-value)	0.085	0.906	0.018	0.028
Panel C: Waiting Time (minutes)				
Line 1 Use	4.53*	5.45***	10.85***	-6.64***
	(2.45)	(1.85)	(4.10)	(1.48)
A-R test (p-value)	0.031	0.039	0.004	0.000
Diagnostic Statistics				
Observations (N)	3,070	4,772	2,126	1,461
K-P Wald rk F-stat	25.31	16.49	17.89	22.94
Hansen J statistic (p-value)	Just-identified	Just-identified	Just-identified	Just-identified
Fixed Effects Included	Yes	Yes	Yes	Yes

Note: Household-clustered standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The models are exactly identified (Just-identified), so the Hansen overidentification test does not apply. Excluded instrument: cumulative proximity ≤ 1 km.

The time-of-day heterogeneity must be interpreted bearing in mind that the composition of trips also varies over the course of the day. The morning peak is dominated by mandatory trips to work or study, whereas evening and nighttime return trips more frequently incorporate the chaining of activities, such as shopping, errands, and care of dependents (McGuckin and Murakami, 1999). Since the proximity instrument depends jointly on the origin and the destination, this asymmetry implies that the effects estimated by time band may reflect both differences in the conditions of the substitute-transport market and changes in the composition of trips and, therefore, in the *complier* population identified in each band.

Consequently, the differences across time bands should be interpreted as joint variations in supply conditions and in trip types, and not exclusively as differences in the metro's return for a given trip type. [4]

[4]The PMU 2023 Mobility Survey records the purpose of each trip, which makes it possible, as a robustness exercise, to re-estimate the model controlling for purpose or restricting the sample to home-based trips for work purposes, with the aim of homogenizing the composition of the trips compared across time bands.

7.4 Distributional Effects and Spatial Sensitivity (NSE)

Since Line 1 operates under a strictly subsidized flat-fare scheme, it is relevant to assess the progressivity of this intervention. Disaggregating the estimate by Socioeconomic Status (NSE) reveals differentiated dynamics that require specialized econometric treatment owing to the heterogeneity in "last-mile" behavior.

Spatial Sensitivity of the Instrument

The baseline instrument with a 1 km radius proves to be an excellent predictor of mode choice for the extreme strata: the Kleibergen-Paap F statistic reaches 50.65 for the High NSE and 11.53 for the Low NSE. Likewise, in these robust specifications, the Anderson-Rubin Wald test firmly rejects the null hypothesis, validating the inference of the coefficients, while the Hansen J statistic remains at 0.000 owing to the exact identification of the model. For the lower-income sectors, strict "walkability" is a relevant determinant; if the station exceeds one kilometer, the additional cost of an approach mode (e.g., motorcycle taxi) dilutes the economic incentive of the metro. Consequently, when the radius is artificially expanded to 2 km for the Low NSE, the relevance of the instrument collapses ($F = 1.41$).

By contrast, for the middle-income stratum (Middle NSE), the 1 km radius turns out to be a weak instrument ($F = 3.44$). This group exhibits a greater willingness to pay for short-distance approach services, functionally expanding the station's area of influence. When the model is calibrated for this segment by expanding the instrument to a cumulative threshold of 2 km ($rank2k_dtotal$), causal identification recovers substantially, raising the F statistic to 8.98. With this calibrated specification, the Anderson-Rubin test again supports the statistical significance of the effect, and the Hansen overidentification test confirms the structural consistency of the single instrument for this population subgroup.

Table 4: Effects on Trip Cost by Socioeconomic Status and Spatial Sensitivity

Variables	(1) High NSE	(2) Middle NSE	(3) Low NSE
Panel A: Baseline Instrument (Proximity $\leq$ 1 km)			
Line 1 Use	-1.33**	-0.34	-2.03***
	(0.54)	(0.61)	(0.76)
Observations (N)	888	1,467	2,454
K-P Wald rk F-stat (1st stage)	50.65	3.44	11.53
A-R Wald test (p-value)	0.017	0.626	0.008
Hansen J statistic (p-value)	Just-identified	Just-identified	Just-identified
Panel B: Expanded Instrument (Proximity $\leq$ 2 km)			
Line 1 Use	-1.82***	-0.67	-6.89
	(0.61)	(0.74)	(5.41)
Observations (N)	888	1,467	2,454
K-P Wald rk F-stat (1st stage)	33.94	8.98	1.41
A-R Wald test (p-value)	0.001	0.395	0.236
Hansen J statistic (p-value)	Just-identified	Just-identified	Just-identified
Global Model Statistics			
Fixed Effects: Origin, Destination, Hour	Yes	Yes	Yes

Note: Dependent variable: total trip cost. Household-clustered standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Panel A uses the strict walkability threshold (1 km), showing instrumental weakness for the Middle NSE. Panel B expands the area of influence to 2 km to capture the last mile of the Middle NSE, improving its identification (F approaches 10), but losing relevance for the Low NSE ($F = 1.41$).

Pro-Poor Impact of the Subsidy

Using the spatially optimal specifications for each stratum (1 km for High/Low NSE; 2 km for Middle NSE), the results confirm the highly progressive character of the infrastructure. The largest monetary saving falls on the most vulnerable populations: Line 1 use reduces trip costs by S/ 2.03 for the Low NSE ($p < 0.01$).

The sign and magnitude of this coefficient are consistent with the fare structure of the system. Since the Line 1 fare is flat and independent of distance, the implicit subsidy per kilometer is increasing in trip length. Lower-socioeconomic-status households systematically make the longest and most peripheral trips (for example, between San Juan de Lurigancho and the central area), whose counterfactual is not a single cheap fare but a chain of chained informal segments—consistent with the larger number of transfers observed in the sample—each priced separately.

Therefore, replacing that chain with a single flat fare generates a larger absolute saving precisely among the lower-income strata. Far from being anomalous, the pro-poor pattern is the expected result of superimposing a subsidized flat fare on a demand in which trip distance and payment fragmentation decrease with income.

By contrast, the High NSE experiences a smaller financial relief, estimated at S/ 1.33 per trip. Finally, the Middle NSE reports the smallest net saving (S/ 0.66, not significant at conventional levels), a result consistent with the spatial sensitivity described above: although the metro is cheaper for them on the trunk fare, part of their gross monetary saving is internally absorbed by the payment of the "last-mile" fare required to connect their origin with the rail infrastructure from distances greater than one kilometer.

It should be stressed that this result refers to the low-income *compliers* with pedestrian access to the network. Since a significant fraction of vulnerable households reside outside the

Line 1 catchment radius, the documented progressivity applies to those who can effectively access the system; the access margin—not the incidence conditional on use—constitutes the main constraint for universalizing these benefits, which reinforces the relevance of the "last-mile" integration policies discussed in Section 9.

7.5 Spatial Robustness Analysis

As a robustness exercise, we assess the sensitivity of the results to the definition of the spatial units used to characterize trips. The main specification uses administrative districts as origin and destination units. Although this definition allows control for spatial heterogeneity through origin and destination fixed effects, administrative aggregation may group together zones with heterogeneous physical and accessibility characteristics. To assess whether the results depend on this spatial delimitation, we estimate an alternative specification using homogeneous spatial units constructed through an H3 hexagonal grid.

In the alternative specification, the route fixed effects are defined through interactions between the origin and destination H3 units, keeping the departure-hour fixed effects and the sociodemographic controls of the main specification. This strategy allows us to compare the estimated effect under a more disaggregated spatial definition, without modifying the instrumental variable or the definition of the treatment.

Table 5: Spatial Robustness of the Line 1 Effect on Trip Speed

	(1) District Model	**(2) H3-R7 Model**
Line 1 Use (LATE)	**13.58*** **	**12.57*** **
	(1.18)	(2.77)
Route fixed effects	Origin + Destination	Origin × Destination
Hour fixed effects	Yes	Yes
Sociodemographic controls	Yes	Yes
KP rk Wald statistic	10.108	17.409
Underidentification LM p-value	0.0583	0.0001
Centered R^2	0.6335	0.8654
Singletons dropped	399	3,150
Observations (N)	**12,285**	**9,534**

Note: Standard errors in parentheses. *** $p < 0.01$. Expansion (survey) weights applied.

The results show that the estimated effect of Line 1 use on speed retains the same sign and statistical significance when the spatial definition is modified. In this exercise, the coefficient on speed goes from 13.58 km/h in the reference district model to 12.57 km/h under the H3-R7 grid; both values are statistically indistinguishable from the aggregate effect of 11.83 km/h (Table 2). The difference in magnitude is moderate, which indicates that the estimated association between Line 1 use and trip speed does not depend exclusively on the district definition of the origin and destination units.

The first stage also provides evidence of instrument relevance in both specifications. The KP rk Wald statistic is 10.108 in the district model and rises to 17.409 under the H3-R7 definition. Likewise, the underidentification LM test rejects the null hypothesis more clearly in the H3-R7 specification. These results provide complementary evidence on the relevance of the instrument under an alternative spatial definition.

Nonetheless, the H3-R7 specification implies a reduction in the sample, from 12,285 to 9,534 observations, associated with the greater granularity of the fixed effects and the elimination of singleton observations. For this reason, this specification is interpreted as a spatial-robustness exercise and not as a substitute for the district model used as the main specification.

7.6 An Approximation to the Economic Value of Travel Time Savings

The estimated causal effects allow a preliminary quantification of the economic value associated with the reduction in travel time generated by the infrastructure. The in-vehicle time saving estimated in the aggregate specification (Table 2), of 27.80 minutes per trip, can be valued monetarily and contrasted with the capital investment in Line 1 (US$ 671.7 million; OSITRAN, 2025).

Considering an annual flow of 191.1 million passengers in 2024 (ATU, 2024), and assuming that each passenger represents one trip for valuation purposes, the aggregate time saving amounts to:

$$27.80\ min \times \frac{1}{60} \times 191{,}100{,}000 \approx 88.5\ million hours per year.$$

To illustrate the economic value of this time saving, we adopt a central value of travel time (VOT) of S/ 5 per hour, together with a sensitivity analysis between S/ 4 and S/ 8 per hour. Under the central value, the annual benefit associated exclusively with the time saving is estimated at approximately S/ 442 million. Under the range considered, this benefit ranges between S/ 354 and S/ 708 million per year.

At an exchange rate of S/ 3.75 per dollar, the capital investment of US$ 671.7 million is equivalent to approximately S/ 2,519 million. The comparison between this investment and the annual benefits derived exclusively from the time saving yields a simple, undiscounted payback period of between 3.6 and 7.1 years, with a central value of 5.7 years. This exercise should be interpreted as an illustrative comparison between CAPEX and the monetary value of the time saving, and not as a financial payback period of the investment.

As a complement, if the annual time-saving benefit is assumed to remain constant over a 30-year useful life and is discounted at a rate of 8%, the present value of the time saving under the central VOT amounts to approximately S/ 4,981 million. Relative to the capital cost of S/ 2,519 million, this implies a ratio of approximately 1.98, that is, close to S/ 2 of present value of time saving for every S/ 1 of initial investment. This comparison incorporates only the benefit associated with travel time and does not constitute a comprehensive cost-benefit evaluation.

These figures should be interpreted as a partial approximation of the economic value generated by the reduction in travel time and not as a complete estimate of the social return to the infrastructure. First, the estimated effect corresponds to a LATE for the *compliers*; its extrapolation to the full set of users therefore assumes that the estimated time saving is reasonably representative of the saving experienced by total demand. Second, the exercise does not incorporate operation and maintenance costs or possible capital-replacement expenditures. Third, it excludes the direct monetary saving for the user (S/ 1.23 per trip), given that, to the extent that this saving reflects a transfer associated with the subsidy scheme, it does not necessarily constitute an additional net social benefit. Finally, other potentially relevant external benefits are also not incorporated, such as the reduction of emissions and pollutants (Morales, 2024) and the alleviation of road congestion. The incorporation of these components, together with operation and maintenance costs, would make it possible to develop a comprehensive social cost-benefit evaluation of Line 1.

8 Conclusions

The present evaluation provides quasi-experimental evidence on the effect of using Line 1 of the Basic Metro Network of Lima and Callao (RBMLC) on the efficiency of urban trips. Through an Instrumental Variables design with High-Dimensional Fixed Effects (IV-HDFE), we use spatial proximity to the stations as a source of exogenous variation to estimate the local average treatment effect (LATE), controlling for individual and trip characteristics, origin and destination fixed effects, and hourly variation. This strategy makes it possible to reduce the concerns associated with mode selection and with unobservable spatial heterogeneity.

The main result shows that Line 1 use increases trip speed by approximately 11.8 km/h, with a positive and statistically significant effect, a magnitude that remains robust (between 12.6 and 13.6 km/h) under alternative spatial definitions. This result indicates a substantial improvement in the temporal efficiency of travel relative to the transport alternatives considered in the sample. In the complementary estimates for the different components of trip efficiency, Line 1 use is also associated with a reduction in in-vehicle time and in trip cost. In particular, the aggregate estimate reports a reduction of approximately 27.8 minutes in in-vehicle time and of S/ 1.23 in trip cost. These results are consistent with the advantage offered by rail infrastructure with an exclusive right-of-way relative to modes subject to the congestion conditions of the road network.

The results also show that the efficiency gains are not independent of the system's operating conditions. The estimated effect on waiting time is positive, which suggests that the benefits associated with a higher travel speed may coexist with higher access and waiting costs. This dimension is relevant for interpreting the effect of Line 1 on the complete travel experience, since a reduction in travel time does not necessarily imply an equivalent reduction in total trip time.

An additional element of the analysis is the assessment of the sensitivity of the results to the spatial definition. The spatial-robustness exercise based on origin and destination districts estimates an effect of 13.58 km/h on speed, close to that of the main aggregate specification. When this spatial structure is replaced by homogeneous H3 units and interacted origin and destination fixed effects are used, the coefficient remains positive and statistically significant, with a magnitude of 12.570 km/h. The closeness between the two estimates provides evidence that the main result does not depend exclusively on the administrative delimitation of districts. This alternative specification is interpreted as a spatial-robustness exercise and not as a substitute for the main model.

Taken together, the evidence indicates that Line 1 generates important improvements in trip efficiency, particularly through a higher travel speed and lower times and costs associated with the trip. At the same time, the existence of higher waiting times and the sensitivity of some results to spatial and operational conditions show that the benefits of the infrastructure must be assessed considering the trip as a whole. These results suggest that future expansions of the mass transit system should complement the expansion of rail capacity with measures aimed at improving accessibility to the stations, reducing waiting times, and managing demand conditions in the main corridors.

Likewise, this study contributes evidence at both the national and international levels, which makes it possible to expand the research agenda developed in studies such as (Contreras et al., 2024) regarding Line 1, as well as the BRT studies (Oviedo et al., 2019; Alba Vivar, 2024; Velásquez, 2023).

9 Policy Recommendations

Based on the econometric evidence generated in this study, the following recommendations are derived, applicable both to the operational management of Line 1 and to the design of the future lines of the RBMLC:

- **Expansion and Subsidy of "Last-Mile" Mobility**
  Since the Middle NSE manages to functionally expand the stations' area of influence to 2 km, but at the cost of eroding its net saving (owing to the payment of motorcycle taxis or shared taxis), it is imperative to formulate policies of physical and fare-based modal integration. We recommend implementing formal feeder-route systems or promoting subsidized micromobility solutions within a radius of 2 to 3 km of the stations. This would make it possible to universalize the system's monetary benefits, attracting more users from the peripheries without penalizing their transport budget.

- **Mitigation of Station Frictions (Peak-Hour Demand Management)**
  The finding of a waiting-time overshoot of 10.9 minutes in the Evening Peak calls for capacity and infrastructure interventions. At the station level, we suggest optimizing entry flows through the full digitalization of fare payment (QR-code reading, contactless bank-card payments) to reduce queues at the turnstiles. At the operational level, we recommend evaluating the insertion of "loop trains" (short services operating only on the most saturated central segment) during demand peaks to inject additional capacity where platform rationing is most acute.

- **Differentiated Fare Policy (*Off-Peak Pricing*)**
  The identified "nighttime anomaly" reveals that, in off-peak-night hours, the train is marginally more expensive (S/ 0.78) than the fragmented informal alternatives, despite its immense time advantages. To maximize the metro's occupancy rate in low-demand hours, the Urban Transport Authority (ATU) should consider implementing differentiated fares (nighttime discounts or *off-peak pricing*). This would not only increase the economic competitiveness of Line 1 relative to informal transport at night, but could also encourage a flattening of the demand curve, partially reducing the pressure in the evening peak. The mechanism behind the "nighttime anomaly"—the disappearance of the congestion surcharge and the competition from fractioned informal fares (Bielich Salazar, 2009; Cervero and Golub, 2007)—implies that a nighttime or *off-peak* discount would not only recover price competitiveness relative to informal transport, but would also make use of the system's idle capacity in low-demand hours.

- **Protection of the Flat Social Fare**
  The distributional results forcefully confirm that the current flat fare (S/ 1.50) functions as a highly progressive mechanism, transferring the largest economic benefit to the Low NSE. We recommend keeping this trunk fare structure unchanged in the face of inflation or operational externalities, financing any deficit through general revenues. Any upward adjustment would have a disproportionate and regressive impact on the city's most vulnerable sectors, forcing them to return to a more informal and more polluting system.

- **Acceleration of Investment in the Integrated Basic Network**
  The magnitude of the time saving (28 minutes) and the increase in speed (11.8 km/h) observed on a single rail corridor are massive by the standard of Latin American megacities. These causal estimates financially justify accelerating the award and construction

of Lines 3 and 4. They also require that the interconnection station with Line 2 (28 de Julio Station) be designed with parametric slack far above the traditional levels, anticipating that network synergy will multiply the spatial *compliers* and place the interchange infrastructure under extreme stress.

References

Abhijna, M., Krishna Rao, K. V., and Perumal, V. (2025). How new metro lines shape a sustainable future: A before-after study of travel behavior, perceptions, and emissions in Mumbai Metropolitan Region, India. *Sustainable Cities and Society*, 124:106431. Published online May 2025.

Acton, B., Le, H. T., and Miller, H. J. (2022). Impacts of bus rapid transit (BRT) on residential property values: A comparative analysis of 11 US BRT systems. *Journal of Transport Geography*, 100:103324.

Ahlfeldt, G. M. and Feddersen, A. (2018). From periphery to core: measuring agglomeration effects using high-speed rail. *Journal of Economic Geography*, 18(2):355–390.

Alba Vivar, F. (2024). Opportunity bound: Transport and access to higher education in a megacity. Doctoral thesis / working paper.

Allport, R. J. (1986). Appropriate mass transit for developing cities. *Transport Reviews*, 6(4):365–384.

ATU (2023). Encuesta de movilidad urbana del PMU 2023.

ATU (2024). Registros administrativos sobre el flujo de pasajeros, periodo 2012–2024.

ATU (2025). Registros administrativos sobre emisiones de la Línea 1, 2015–2024.

Autoridad de Transporte Urbano para Lima y Callao (ATU) (2025). Plan de movilidad urbana (PMU) 2025–2045 para lima y callao. Nota de prensa, Plataforma del Estado Peruano. Disponible en `https://www.atumovilidad.pe`.

Ayaragarnchanakul, E., Klauber, H., Nachtigall, F., Wagner, F., Hu, J.-w., and Creutzig, F. (2025). The impact of the Bangkok metro expansion via big taxi GPS probe data. *npj Sustainable Mobility and Transport*, 2(1):4.

Banco Central de Reserva del Perú (2024). Reporte de inflación. diciembre 2024. Technical report, BCRP.

Bielich Salazar, C. (2009). La guerra del centavo: Una mirada actual al transporte público en lima metropolitana. Documento de Trabajo 155, Instituto de Estudios Peruanos (IEP) y Consorcio de Investigación Económica y Social (CIES), Lima.

Bocarejo, J. P., Portilla, I. J., Velásquez, J. M., Cruz, M. N., Peña, A., and Oviedo, D. R. (2014). An innovative transit system and its impact on low-income users: the case of the Metrocable in Medellín. *Journal of Transport Geography*, 39:49–61.

Brinkman, J. and Lin, J. (2026). The costs of access: Freeway disamenities across us cities. *Regional Science and Urban Economics*, page 104253.

Cervero, R. and Golub, A. (2007). Informal transport: A global perspective. *Transport Policy*, 14(6):445–457.

Céspedes-Reynaga, N. (2025). El muy prolongado viaje al trabajo en el Perú. Documento de Trabajo 2025-009, Banco Central de Reserva del Perú, Lima.

Cheng, J. and Chen, Z. (2021). Socioeconomic impact assessments of high-speed rail: A meta-analysis. *Transport Reviews*, 42(4):467–502.

Comisión Investigadora del Caso Lava Jato (2018). Informe final de la comisión investigadora del congreso de la república sobre el caso Lava Jato en el Perú. Technical report, Congreso de la República del Perú.

Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly exogenous. *The Review of Economics and Statistics*, 94(1):260–272.

Contreras, C., Aguilar, D., and Lázaro, J. (2024). Impacto del transporte masivo en la informalidad de vivienda: El caso de la Línea 1 del Metro de Lima. *Argumentos*, 5(1).

Corporación Andina de Fomento (CAF) (2015). El metro de lima: Impacto en la movilidad y accesibilidad urbana. Technical report, CAF.

Daniels, R. and Mulley, C. (2013). Explaining walking distance to public transport: The dominance of public transport supply. *Journal of Transport and Land Use*, 6(2):5–20.

de Sá, A. L. S., Lavieri, P. S., Pawlak, J., Sivakumar, A., and Thompson, R. G. (2025). The effects of travel time use on activity-travel behaviour: knowledge consolidation and research agenda for current and future transport options. *Transport Reviews*, pages 1–28.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica*, 78(6):1863–1903.

Diao, M., Leonard, D., and Sing, T. F. (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics*, 67:64–77.

Donaldson, D. (2025). Transport infrastructure and policy evaluation. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 6. Elsevier.

Dong, H. (2017). Rail-transit-induced gentrification and the affordability paradox of TOD. *Journal of Transport Geography*, 63:1–10.

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

Giannopoulos, G. A. (1981). Metropolitan railways: present characteristics and future prospects. *Transport Reviews*, 1(1):45–74.

Gonzalez-Navarro, M. and Turner, M. A. (2018). Subways and urban growth: Evidence from earth. *Journal of Urban Economics*, 108:85–106.

Graham, D. J. (2025). Causal inference for transport research. *Transportation Research Part A: Policy and Practice*, 192:104324.

Guerrero, T. E., Guevara, C. A., Cherchi, E., and Ortúzar, J. d. D. (2021). Addressing endogeneity in strategic urban mode choice models. *Transportation*, 48(4):2081–2102.

GyM Ferrovías (2014). Reporte de sostenibilidad 2013–2014. Technical report, GyM Ferrovías / Línea 1.

GyM Ferrovías (2018). Reporte de sostenibilidad 2018. Technical report, GyM Ferrovías / Línea 1.

GyM Ferrovías (2024). Mapa de estaciones – Línea 1 del Metro de Lima.

Hu, J., Lu, H., Li, M., Long, Y., Li, Y., and Zeng, Y. (2024). Impact of newly extended metro line opening on commuting carbon emissions: A comparative study before and during the pandemic. *Transport Economics and Management*, 2:203–213.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Instituto Nacional de Estadística e Informática (2025). Encuesta nacional de uso del tiempo (ENUT) 2024. Microdatos y ficha técnica.

JICA (2012). Encuesta de recolección de información básica del transporte urbano en el área metropolitana de Lima y Callao. Technical report, Japan International Cooperation Agency.

Keele, L. J. and Titiunik, R. (2015). Geographic boundaries as regression discontinuities. *Political Analysis*, 23(1):127–155.

Kim, K. et al. (2021). Impacts of light rail transit on labor participation and housing affordability in the u.s.: Longitudinal analysis using propensity score matching. *Journal of the Transportation Research Board*. https://doi.org/10.1177/03611981211029650.

Lee, D. S., McCrary, J., Moreira, M. J., and Porter, J. (2022). Valid t-ratio inference for IV. *American Economic Review*, 112(10):3260–3290.

Llosa, L. G. and Panizza, U. (2015). La gran depresión de la economía peruana: ¿una tormenta perfecta? *Revista Estudios Económicos*, 30:91–117.

Matas, A., Raymond, J. L., and Roig, J. L. (2020). Evaluating the impacts of HSR stations on the creation of firms. *Transport Policy*, 99:396–404.

McGuckin, N. and Murakami, E. (1999). Examining trip-chaining behavior: Comparison of travel by men and women. *Transportation Research Record*, (1693):79–85.

Ministerio de Economía y Finanzas (MEF) (2025). Seguimiento de proyectos de inversión: Líneas 3 y 4 de la RBMLC. https://www.gob.pe/institucion/mtc/noticias/1063952-mtc-presento-proyectos-de-las-lineas-3-y-4-del-metro-de-lima-y-callao.

Ministerio de Transportes y Comunicaciones (MTC) (2024). Avances en la implementación de la Línea 2 del Metro de Lima y Callao.

Ministerio de Transportes y Comunicaciones (MTC) (2026). Línea 2 del metro de lima y callao supera el 83 % de avance físico de obra y transformará la movilidad urbana. Nota de prensa, Plataforma del Estado Peruano. Disponible en `https://www.gob.pe/institucion/mtc/noticias/1395210`.

Mohammadi, A., Amador-Jimenez, L., and Nasiri, F. (2019). Review of asset management for metro systems: challenges and opportunities. *Transport Reviews*, 39(3):309–326.

Morales, G. (2024). Beneficios ambientales del transporte férroviario eléctrico: Evidencia de la Línea 1 del Metro de Lima. Tesis de Maestría en Desarrollo Ambiental. Pontificia Universidad Católica del Perú.

OSITRAN (2024). Informe de desempeño: Línea 1 del Metro de Lima y Callao.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Oviedo, D., Meléndez Fuentes, N., and Chong, A. Z. W. (2025). Transport affordability across time and space: Considerations for Latin America and the Caribbean and other Global South settings. *Journal of Planning Literature*. Published online 2025; advance online publication.

Oviedo, D., Scholl, L., Innao, M., and Pedraza, L. (2019). Do bus rapid transit systems improve accessibility to job opportunities for the poor? the case of lima, peru. *Sustainability*, 11(10):2795.

Redding, S. J. and Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5, pages 1339–1398. Elsevier.

Scholl, L., Martinez, D., Mitnik, O. A., Oviedo, D., and Yañez-Pagans, P. (2018). A rapid road to employment? the impacts of a bus rapid transit system in lima. Technical report, IDB Working Paper Series.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Andrews, D. W. K. and Stock, J. H., editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108. Cambridge University Press, Cambridge.

Tan, Y., Zhao, P., and Li, L. (2025). Subway expansion, residential relocation, and travel behavior: Causal evidence from China. *Journal of Transport Geography*, 124:104166.

TomTom Traffic Index (TTI) (2024). Traffic index 2024: Lima.

Tsivanidis, N. (2026). Evaluating the impact of urban transit infrastructure: Evidence from Bogotá's transmilenio. *American Economic Review*, 116(2):418–463.

Välilä, T. (2025). The economic impact of transport infrastructure: a review of project-level vs. aggregate-level evidence. *Transport Reviews*, 45(4):459–481.

Van den Berg, R. and De Langen, P. W. (2017). Environmental sustainability in container transport: the attitudes of shippers and forwarders. *International Journal of Logistics Research and Applications*, 20(2):146–162.

van Kippersluis, H. and Rietveld, C. A. (2018). Beyond plausibly exogenous. *The Econometrics Journal*, 21(3):316–331.

Velásquez, D. (2023). Transit infrastructure, couples' commuting choices, and gender earnings inequality. *Unpublished manuscript*.

Wang, C., Meng, W., and Hou, X. (2020). The impact of high-speed rails on urban economy: An investigation using night lighting data of chinese cities. *Research in Transportation Economics*, 80:100819.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2 edition.

Wu, W., Zheng, S., Wang, B., and Du, M. (2020). Impacts of rail transit access on land and housing values in china: a quantitative synthesis. *Transport Reviews*, 40(5):629–645.

Yang, H., Tong, Q., and Zhao, P. (2022). Study on the impact of the Beijing-Shanghai high-speed railway on regional economic equity based on economic performance. *The Annals of Regional Science*, 68(1):229–253.

Zhang, S., Lo, H. K., Ng, K. F., and Chen, G. (2021). Metro system disruption management and substitute bus service: a systematic review and future directions. *Transport Reviews*, 41(2):230–251.