Transit Infrastructure and Couples' Commuting Choices in General Equilibrium

Daniel Velásquez*
University of Michigan

First version: August 2021
This version: January 2023
Newest version here

Abstract: What is the impact of improving the transit infrastructure on the gender earnings gap? How does family structure matter to understand the impact of new transit infrastructure? Recent models on spatial economics hinge on the assumption that households are comprised of a single type of person making commuting and location choices. In reality, an important share of the population lives in households with multiple members, whose commuting choices might be interlinked through the household's budget constraint. In particular, as the wage of one of the spouses increases and secures a higher standard of living, households are willing to give up more of the other spouses' income to decrease commute costs. This way, reduced commute times impact one partner's commuting patterns not only by affecting her prospects, but also those of her spouse. I set up and estimate a quantitative model of city structure featuring single and married households leveraging on the introduction of a Metro line and the Bus Rapid Transit System (BRT) in Lima, Peru. My model delivers interdependent commuting choices within dual-earner households. Despite this complexity, commuting follows a gravity structure which provides tractability and greatly simplifies estimation. In the counterfactual analysis, I find that the gender gap in aggregate earnings decreased by over 13% in remote areas where access to improved infrastructure increased the most. Conclusions would change substantially if I use a model with independent commuting choices.

Keywords: Mass Transit, Gender Gap, Family, Dual-Earners, Quantitative Spatial Models, Latin America.

---

*Department of Economics, University of Michigan (email: danielvc@umich.edu). I am very grateful to Maria Aristizabal-Ramirez, Luis Baldomero, Dominick Bartelme, Agostina Brinatti, Alberto Chong, Don Davis, Jorge De la Roca, Luis Espinoza, Andrei Levchenko, Luis Quintero, Ana Reynoso, Chris Severen, Sebastian Sotelo, Dean Yang, and Basit Zafar for very useful comments and suggestions. I also thank Martín Carbajal for excellent research assistance. I am also grateful to CAF-Development Bank of Latin America for financial support. The standard disclaimer applies.

1 Introduction

Does family structure play a role in understanding the impact of new transit infrastructure? What is the impact of improving commute times on the gender earnings gap? As urban populations grow in developing countries, policymakers are becoming increasingly concerned about the strain on transportation systems and the potential for exacerbating inequalities in access to opportunities. For example, previous research has shown women value commuting more than men (Le Barbachon, Rathelot, and Roulet, 2021). Thus, investing in the transit system may be a logical response to address these concerns.

Nonetheless, current spatial models are not fully equipped to study these issues. These models rely on the assumption that households consist of a single person making individual choices about consumption, housing, commuting, and location (Ahlfeldt, Redding, Sturm, and Wolf, 2015; Heblich, Redding, and Sturm, 2020; Tsivanidis, 2021). In reality, a significant portion of the population lives in households with multiple members, where incentives to commute are governed by different considerations than in single-person households. For example, Figure 1 shows that married individuals in urban Peru account for more than 50% of all hours spent commuting. This figure also reveals a marked specialization among married individuals, with males in married households accounting for over 70% of commuting time, compared to roughly 50% for single males. This pattern is also true for dual-earner households, as males account for about 60% of all commuting time.

I build upon previous spatial models by examining whether the composition of households plays a role in understanding the impact of new transit infrastructure on the gender earnings gap. Through my analysis, I uncover a new mechanism: that commuting choices are interlinked through the household's budget constraint. In particular, as the wage of one of the spouses increases and secures a higher standard of living, households are willing to give up more of the other spouses' income to decrease commute costs. Reduced commute times can therefore affect commuting choices through two channels: by affecting an individual's prospects, and by affecting those of their spouse. This has not been considered in previous literature. Additionally, by considering married households, I am able to study how specialization and the interlinked nature of commuting choices may interact with location and specialization choices.[1]

[Figure 1 here]

I focus on two infrastructure investments that were introduced around 2011 in Lima Peru: the Line 1 of the Metro and the Bus Rapid Transit system (BRT). These infrastructure projects have contributed to shaping the city's landscape by connecting the fringe of the city to central locations and improving commuting times by around 25-45%. Moreover, Lima is an ideal setting since it might be representative of other cities in the developing world. It is a very dense city with a population of more than 10 million and substantial transit congestion.

[1]For example, If a couple shares a budget and one partner has a high-paying job in the central business district (CBD) while the other has a lower-paying job farther from the CBD, they may choose to live farther from the CBD to save on housing costs. On the other hand, if both partners have high-paying jobs in the CBD, they may be willing to pay a premium for housing in the CBD in order to reduce commuting costs.

I assemble a data set consisting of population censuses (1993, 2007, and 2017), an economic census (2008), a cadaster of firms (2015), rich survey data (2007-2018), and road network data, and document two empirical facts suggesting that Lima's Metro and BRT may have had not only gendered impacts, but that family structure might matter. First, commuting elasticities are heterogeneous across households. Women's commuting flows are more sensitive to improvements in commuting times than men's, especially when women cohabit with a partner. Second, the choice of where to work depends on the spouse's commuting choice, suggesting that choices are interdependent within dual-earner households. I complement these facts with a set of reduced-form impacts. I leverage the planned-but-not-constructed lines to build a fitting control group. This way, I compare locations close to Metro or BRT stations, against locations close to planned lines, before and after the Metro and the BRT started operating.

It is hard to extract definitive conclusions from this analysis since the introduction of these transit investments potentially affected the whole city, blurring the lines between what control and treatment groups are. To better understand these issues and tease out the different forces that are interacting, I, first, develop a partial equilibrium model of commuting for married households where commuting is costly. Then, I embed this model into a spatial quantitative model of city structure.

I microfound the partial equilibrium model by utilizing a standard framework in family economics, that is, by assuming that married households achieve Pareto efficiency. Two main insights are derived from this application.

The first insight is that commuting choices are interdependent across spouses. Hence, there are two channels through which improved commute times can affect households. One is the direct use channel, and the other one is the indirect use channel (or within-household spillovers). The direct use channel is the mechanism present in the standard class of spatial models. It says that when times between origins and destinations improve, income at origins increases only if the new destinations are high-wage locations. Otherwise, if the new destinations are low-wage locations, improvements in commuting times may induce some workers to give up some of their earnings to have a shorter commute. This means that there is potentially a trade-off between wages and commuting. My contribution to the literature is the indirect use channel. This channel says that when commute costs for one of the spouses decreases, then this affects the households' choice of the other spouse's destination. To see why, note that under this framework, the household dislikes the commuting performed by the two members. Hence, when one spouse improves his or her commuting times, this affects the trade-off between wages and commuting times for the other spouse, leading the household to potentially chose a new destination for the second spouse.

The second insight is that commuting flows follow a gravity structure which makes the model tractable and easy to estimate. Thus, my contribution is to set up a model delivering commuting gravity equations, even for married households with interdependent commute choices.

In the quantitative section of the paper, I embed the model of commuting choices for married households into a general equilibrium model of city structure based on Ahlfeldt et al. (2015) and Tsivanidis (2021). This model includes important elements from the literature, as well as features specific to the setting of this study. For example, the model considers both single and married households, and allows for commuting elasticities to be heterogeneous across genders and civil status. It also includes a residential location choice with location elasticities that differ between single and married households. Married households can also choose to be either dual-earner or male-breadwinner households, meaning that they can evaluate whether to send both spouses to the labor market or to have one spouse specialize in household production. Thus, in practice this is a labor participation choice for married women.[2] Additionally, from the production side, the model includes multiple sectors that employ workers of each gender with a certain degree of substitution, allowing for the presence of gender gaps in the workplace.[3] Finally, the model allows for endogenous sectoral TFP to account for density effects (Ahlfeldt, et al., 2015; Tsivanidis, 2021). These elements are crucial in generating the gendered impacts of new transit infrastructure.

I calibrate and estimate the model using Peruvian data. To estimate commuting elasticities, I exploit data on bilateral flows and the gravity structure of the model for commuting probabilities. Moreover, to estimate the remaining parameters, I leverage on Lima's road network and the introduction of the new transit infrastructure to construct Bartik-style instruments. Once parameters are estimated I show that the model with interdependent commuting choices within couples fits better the data on bilateral commuting flows of dual-earner households than the model with independent commuting choices.

I simulate what Lima would look like in the absence of the Metro and the BRT under interdependent commuting choices, and under a model with independent commuting choices. First, using the model with interdependent commuting choices I find that the Line 1 of the Metro and the BRT increased aggregate real income by 1.19%. Moreover, the gender gap in aggregate earnings decreased by 2.07%. This is mainly driven by married couples. The gender gap in aggregate earnings decreased by 4.00% for married households, but increased by 0.49% for singles. This hides substantial heterogeneity. For instance, the gender gap among couples decreased by more than 13% in the fringe of the city, that is, in the set of locations that gained access the most. I then use the model to rationalize the set of reduced-form results. Finally, I use the model to decompose the margins through which better commuting times affected households. This exercise reveals that the new transit infrastructure contributed the most in reducing gaps through the direct use channel and the labor force participation choice. However, the indirect use channel and other general equilibrium effects tended to widen the gaps.

Results are substantially different if I use a model with independent choices instead. First, while this model also suggests that the gap in aggregate earnings decreased for married households, it underestimates the decrease by more than 40%. Second, according to this model, the direct use channel did not have a fundamental role in reducing the gap. There is no reason why the direct use channel inferred by the two models should be the same. The direct use channel depends on the geography of commute probabilities, and commute time changes. Whereas commute time changes are the same in the two models as they are data, the probabilities were backed out thanks to the

[2] I assume that married households are comprised of a male individual, i.e. the husband, and a female individual, i.e. the wife. A simplifying assumption usually done in the literature on family economics is that the husband always works. Moreover, in the data female breadwinner households constitute just 1.3% of all households. This is why I turn off this margin: it does not seem to be quantitatively important but would add greater complexity to the model.

[3] For instance, see Acemoglu et al. (2004).

models' structures and the observed data. Without precise data on commute flows, the model with independent choices would not have been able to adequately measure married women's direct exposure to the new infrastructure.

Last but not least, I perform a counterfactual exercise to predict what would happen if the planned network of Metro lines is constructed. My results show that real earnings in the city would increase by an another 1%, and that the gender gap would decrease by an additional 1.9%. This means that the complete Metro network and the BRT would have increased real earnings by a total of 2.22%, and decreased the gender gap by a total of 3.9%. Furthermore, the gap would decrease by a total of 7.9% for married households, and would increase by 1.9% for singles. Once more, these are aggregate impacts and hide substantial heterogeneity as not all locations are close to new stations.

Literature

This paper is related to many branches of the literature. First, my work speaks to the recent literature on economic geography models,[4] and in particular of cities.[5] This literature examines the impact of transit infrastructure on economic outcomes accounting for spillovers through the transit network in a general equilibrium framework. I incorporate a new mechanism positing that commute choices are interdependent within dual-earner households. Thus, reduced commute times can affect commuting choices through two channels: by affecting an individual's prospects, and by affecting those of their spouse. I set up a model that preservers the gravity structure of commuting while allowing for this interconnection. Ignoring interdependence in commuting choices changes conclusions substantially.

Second, this paper relates to the literature on family economics. There are two major strands: what happens inside existing unions and who marries whom (Browning et al., 2011).[6] My work mainly contributes to the first strand by showcasing how relaxing time constraints can impact earnings inequality within households through spouses' choices. This paper also illustrates how the family dynamic can influence decisions related to the spatial aspect of the model, such as choices about where to live and where to commute.

Third, this paper relates to the literature that examines gender differences in urban labor markets. This literature documents that there are gender differences in location choices and commuting choices. These gender differences tend to intensify among the married. For instance, Black et al. (2014) document a very large variation in the labor supply of married women across US cities due to differences in commuting times.[7] Moreover, recent experimental evidence reveals that reducing mobility constraints has a large impact on job searching for women (Field and Vyborny, 2022). Additionally, there is a huge

[4] See Allen and Arkolakis (2014), Behrens et al. (2017), Bartelme (2018), Monte et al. (2018), Bryan and Morten (2019), Allen and Arkolakis (2020), Fajgelbaum and Schaal, (2020), Adao et al. (2020), Kwon (2020).

[5] See Ahlfeldt et al. (2015), Allen and Arkolakis (2015), Tsivanidis (2021), Owens et al. (2020), Heblich et al. (2020), Severen (2021), (Zarate) 2021, Khanna, et al. (2022).

[6] See Chiappori, et al. (2022).

[7] Kawabata and Abe (2018) examine the relationship between commute times and female labor force participation within Tokyo, showing that it differs markedly across household types. Mehmood et al. (2021) document similar findings for India. Carta and De Philippis (2018) explore the relationship between husbands' commuting times and wives' employment and family time allocation. Their findings reveal that a 1% increase in the husband's commuting distance reduces his wife's employment probability by 0.016. Other relevant papers include Clark et al. (2002), Rosenthal and Strange (2012), Gutierrez (2018), Farré et al. (2020), Kwon (2020), and Gu et al. (2021).

literature on the trade-off between wages and non-wage job attributes. For example, Le Barbanchon et al. (2021) show that women value commute 20% more than men.[8] My contribution to this literature is to incorporate the family and spatial dimension together into a single tractable framework, allowing me to quantify the aggregate impact of new transit infrastructure. As a by-product, this framework also provides me with the means to quantify the importance of female labor force participation in shaping the aggregate impact of transit infrastructure.

Fourth, recent work in macroeconomics suggests that family and decision-making in families should be an integral part of economic models since the family structure has repercussions on the aggregate labor supply (Doepke and Tertlit, 2016).[9] My work contributes to this debate by showcasing that, even in a static model, the family structure may have important implications on the determination of aggregates once that we account for the spatial distribution of economic activity.

Finally, this research contributes to the literature on the estimation of agglomeration effects and their quantitative importance. In particular, most estimates of agglomeration forces exploit variation of *overall* employment *across regions* for identification.[10] Rather I estimate *sectoral* agglomeration forces *within a city*.[11]

Paper's organization

The paper is organized as follows. Section 2 briefly describes the Metro and the Bus Rapid Transit system. Section 3 describes my data sources and shows three stylized facts. Section 4 develops the a partial equilibrium model of commuting in married households. Section 5 embeds this partial equilibrium model into a quantitative spatial model, and explains choices regarding calibration and estimation of the model's parameters. Section 6 removes the BRT and the Metro from the city to evaluate their impact. Section 7 performs a counterfactual analysis introducing planned-but-not-yet constructed Metro lines. Section 8 concludes.

2 Lima's Metro and Bus Rapid Transit System (BRT)

Both transit investments were introduced around 2011 in Lima Peru, and since then, they have contributed to shaping the city's landscape. The Metro is 34.4 km long and serves more than 500,000 passengers every day. The first 22 km of the Metro were finished in 2011, while the remaining 12.4 km were finished in 2014. Although the Metro does not cross through the CBD, it crosses through high-demand areas, and in particular, through one of the most important commercial areas in the city, that is, the market of "Gamarra" (see Figure B.1). Moreover, thanks to the Metro, commuting

[8]See Becker (1985), Cha and Weeden (2014), Goldin (2014), Angelov et al. (2016), Lundborg et al. (2017), Adda et al. (2017), Reuben, Wiswall, and Zafar (2017), Cubas et al. (2019), Kleven et al. (2019), Mas and Pallais (2020), Liu and Su (2020).

[9]See also Mankart and Oikonomou (2017), Greenwood et al. (2021), Alon et al. (2022), Bardóczy (2022).

[10]Rosenthal and Strange (2004), Combes et al. (2008), Greenstone, Hornbeck, and Moretti (2010), De la Roca and Puga (2017).

[11]There are only a few papers estimating sectoral agglomeration forces within a city (e.g. Zarate, 2021). Examples of estimates within city for a single sector are Ahlfeldt et al. (2015), Tsivanidis, (2021), and Heblich et al. (2020).

times were greatly improved. For example, from one end to the other, commuting times were reduced from 2 hours and 45 minutes to just an hour.

The BRT was introduced around 2010. Typically, a BRT system is comprised of a system of roadways that are dedicated to buses. They have been successfully implemented in cities like Bogota and Buenos Aires. As it can be seen in Figure B.2, Lima's BRT system connects the fringe to the modern districts, in addition to the central business district (CBD). It is 26 km long, serving around 700,000 passengers daily. The idea of a BRT system is to provide better speeds while keeping the simplicity of a system of buses. In line with this, a study by the World Bank reports that commuting times were cut by around 25-45% (Oviedo et al., 2019).

3 Data Sources, Stylized Facts, and Reduced-form Evidence

I start by briefly summarizing the data sources. Then, I explain two stylized facts highlighting how commuting might be different for couples compared to singles. Last but not least, I present a set of reduced-form impacts of the new transit infrastructure motivating the quantitative exercise.

3.1 Data sources

For this paper's analysis, I require several sources of data geographically identified at a fine scale. A zone is a set of 25 blocks approximately. A block is the finest available disaggregation used by the National Institute of Statistics and Informatics. Lima in 2017 was partitioned into more than 30,000 blocks and more than 1500 zones. My analysis is performed always at the zone level. Hence, throughout the paper, I refer to zones and blocks interchangeably.[12]

In particular, (i) I use the Population Census from 1993, 2007 and 2017 to get residence counts at the zone level for single men and women, and cohabiting couples. (ii) I also use the Economic Census of 2008 and an open administrative data for firms' location and employment figures in 2015 to calculate employment levels at the block-by-industry level (available here). To compute commuting flows, I rely on several sources of data. (iii) I use the Population Census from 2017, which has data on commuting flows at the municipality level. (iv) I complement this with several commuting surveys collected during 2010-2018 for the city of Lima. Then, (v) I use data on the road network from Open Street Maps and Google Maps API. Furthermore, (vi) I use the National Household Survey (ENAHO) from 2007-2017 waves, which is precisely georeferenced, to analyze rents at the zone level. (vii) Data on land use come from the Metropolitan Institute of Planning of Lima. Finally, (viii) I use the 2010 National Time Use survey to complement the data on commute times. Using my data set, I document two stylized facts that motivate my empirical design and modeling approach.

More information about the data is available in the Data Appendix, which can be found in Section A1.

[12] I also use the word "location" to refer to zones. When talking about commuting, I use the word "origin" to refer to the zones where households live, and "destination" to refer to the zones where households work.

3.2 Stylized facts

The following stylized facts motivate the idea of incorporating married households into standard spatial models.

Fact 1: Commuting elasticity differs by civil status and gender. Using the Population Census, I classify households into five different groups. The first two consist of (i) *single females* and (ii) *single males*, which comprises 31.9% and 35.9% of households in 2017. Then I consider (iii) *male breadwinner* households, that is, households where the wife stays at home while the husband goes to work (13.3% of households). For the sake of completeness, I also consider (iv) *female breadwinner* households, although they just constitute 1.3% of households. Finally, I consider (v) *dual-earner* households, which represent 17.6% of households. For each of these groups, I run the following reduced form regression:

$$\log \pi_{j|i,k} = \beta_k time_{ij} + FE_{i,k} + FE_{j,k} + \varepsilon_{ij} \tag{1}$$

where $\pi_{j|i,k}$ is the share of members of civil status-by-gender k that live in origin location i that commute to location j.[13] $time_{ij}$ is the time that takes to commute from i to j. Moreover, $FE_{i,k}$ and $FE_{j,k}$ are fixed effects by origin i and destination j. Finally, β_k is the *commuting elasticity*, that is, it measures how the probability of commuting to j would change in relative terms after increasing the commuting time by one minute.

To compute commuting probabilities $\pi_{j|i,k}$ I rely on the 2017 population Census, and aggregate the data at the district level.[14] I measure commuting times by zones using two data sources: Open Street Maps (OSM) and the 2010-2018 waves of Lima's Commuting Survey. I assign speeds to match documented speed times in Lima. I then aggregate these times to the district level by taking the median value across zones within districts. I show that these times are well correlated with times reported in commuting surveys (see Figure B.3).

Table 1 shows the result of estimating this equation by PPML. Columns (1) and (2) show commuting elasticities for single males and females, respectively. These estimates suggest that single females are 20% more sensitive to changes in commuting times than single males. This difference is statistically significant at the 1% level. Columns (3) and (4) exhibit these estimates for breadwinner households. Comparing males and females in breadwinner households leads to a similar conclusion. Females are more sensitive than males in breadwinner households and the gap widens to 36%. This difference is statistically significant at the 1% level. Finally, columns (5) and (6) deliver the estimates for males and females in dual-earner households, showing that females are 40% more sensitive than males. This difference is statistically significant at the 1% level.

Figure 2 depicts the bin scatter between the residual of log commuting probabilities, net of origin and destination fixed effects, and the residual of commuting times. Panel A does it for singles, while Panel B for married in male and female breadwinner households. The graph confirms the log-linear

[13] I use the word 'Married' to refer to couples in a general sense. See Section A1 for more information.

[14] While I can locate the origin of households at the block level, the data only records the destination district, and so I perform this analysis at the district level.

relationship between commuting probabilities and commuting times, and the fact that the gap in elasticities is wider among the married.

All in all, these results show that women are more sensitive to improvements in commuting times than men especially when cohabiting with a partner, which is in line with previous findings in the literature about the trade-off between wages and non-wage job attributes.[15] Hence, the model should allow for heterogeneity in commuting elasticities.

[Table 1 here]

[Figure 2 here]

Fact 2: Choice of where to work depends on spouse's choice. In dual-earner households, the choice of where to work may depend on the spouse's destination choice. To see this, first, note that if the household cares about some weighted average of the members' utility, then the household, will dislike members commuting to locations farther away, conditional on total income. Second, conditioning on the first spouse's destination, the utility that the household derives from sending the other spouse to some destination may depend on the income the first spouse is already bringing to the household. For example, if one of the spouses is already commuting to a high wage location, thus increasing household income, then the other spouse's supply may be less sensitive to further increases in her own wage. This is akin to a classic result in labor economics where an income effect may reduce the labor supply elasticity (Blundell et al., 2007). In this paper, rather than focusing on an intensive margin defined over the number of working hours, I study the distance one commutes and so a similar logic applies.[16]

I can test whether commuting probabilities of one spouse depend on those of the other spouse by leveraging on the fact that I observe each member's commuting choice in dual-earner households. This allows me to compute *conditional* commuting probabilities. In turn, this permits me to test whether the interaction of destinations is predictive of conditional commuting probabilities once that I partial out each member's independent reasons to work in such locations, such as expected earnings. In a world with interdependence, the interaction of destinations should be highly predictive of conditional commuting probabilities once that I control for each members' independent reasons to work in a particular destination pair. For example, a particular combination of wages at a destination pair may provide greater utility to the household than other combinations, given commute costs.

I denote the conditional probability of member g of working in location j given that the spouse commutes to j' as $\pi_{j|ij',g}$. Thus, if the city has say 50 locations, this object would have $50 \times 50 \times 50$ cells as one spouse living in i can commute to j, while the other spouse can commute to j'. I then estimate the following regression for each member in dual-earner households:

$$\log \pi_{j|ij',g} = \beta time_{ij} + FE_{i,j',g} + FE_{j,j',g} + \varepsilon_{ijj',g} \tag{2}$$

where $FE_{i,j',g}$ are a set of origin-spouse's destination fixed effects, which control for the fact that certain origin-spouse's destination pairs may provide greater access to labor markets to member g.

[15] See Becker, 1985; Cha and Weeden, 2014; Goldin, 2014; Angelov et al., 2016; Lundborg et al., 2017; Adda et al., 2017; Reuben, Wiswall and Zafar, 2017; Cubas et al., 2019; Kleven et al., 2019; Mas and Pallais, 2020.

[16] In this paper I focus on the income channel within couples rather than scale economies when commuting.

Special focus should be given to $FE_{j,j',g}$, which are a set of destination pairs fixed effects controlling for the fact that certain destination pairs might be more desirable than others.

My procedure goes as follows. First, I estimate the previous equation. Then, I recover the estimates of $FE_{j,j',g}$. If one's choice depends on the other spouse's choice, then the residual of $FE_{j,j',g}$—after controlling for fixed effects at one's own location j and the spouse's location j'—should be predictive of commuting probabilities as noted above.[17] That is, the coefficient ψ from $\log \pi_{j|j',g} = \psi_g \hat{FE}_{j,j'} + FE_j + FE_{j'}$ should be different than zero. The greater the value of ψ, the lesser that choices are independent.[18]

I show that this is in fact true. Figure 3 depicts a scatter plot between residualized commuting probabilities and the residuals of $FE_{j,j'}$. Panel A depicts this plot for husbands while Panel B, for wives. We can see that residualized $FE_{j,j'}$ are highly predictive of residualized commuting probabilities for both spouses. In fact, $\hat{\psi} = 0.87$ for males, while $\hat{\psi} = 0.54$ for females.

[Figure 3 here]

One could still argue that this procedure is capturing anything that varies at the destination-pair level, but that is unrelated to the joint decision that partners make regarding their commuting. I now provide suggestive evidence of the income effect explained above. As such, I lose the destination-pair fixed effect in equation 2 and instead add the average wage at destination j (where g commutes) and the average wage at destination j' (where the spouse commutes). Moreover, I interact wages at one's destination with the total income they would obtain at a destination pair.[19] I show the results of this specification in Table B.1. These results show that the greater the household income, the lesser the sensitivity to further increases in one's own wage. Moreover, there is some heterogeneity around the average own-gender wage elasticity.[20] In the case of husbands in dual-earner households, this elasticity goes between 1.856% and 3.054% (column 1), whereas for wives this elasticity goes between 1.178% and 1.914% (Column 2).

To sum up, the evidence I provide suggests that commuting choices are interlinked. In the theoretical and quantitative section of the paper I quantify how important is this interdependence when evaluating the impact of new transit infrastructure.

3.3 Reduced-form evidence

As it will become clear in the theoretical section of the paper, the new transit infrastructure can impact households through many channels. The main channel is the improvement in commuting times leading to greater access to jobs. However, the Metro and BRT can also affect local exogenous amenities, for instance by bringing more crime or noise, and the price of the land (Brinkman and Lin, 2022). In this section, I explore these issues in the context of reduced-form specifications.

[17] I also control for the time it takes to go from workplace in j to workplace in j' to take into account that couples may decide destinations partially based on the distance between jobs.

[18] For simplicity note that I have averaged $\pi_{ij|j',g,k}$ over origins, and so $\pi_{j|j',g}$ is the average probability than one commutes to j given that his or her spouse commutes to j'.

[19] I obtain average wages for each gender from the 2015-2017 waves of the national firm-level survey data.

[20] I have demeaned the variable for log total income, hence the coefficient associated to one's own gender's wages reveals the average elasticity.

3.3.1 Planned stations

I perform a reduced-form analysis by exploiting the planned network of Metro lines. Specifically, I leverage Line 2 and Line 4 of the Metro.[21] I compare locations closer to Line 1 of the Metro or the BRT (treatment locations) against locations closer to Line 2 and Line 4 (control locations). The basic idea is that we can account for location selection effects by comparing treated locations to locations closer to planned-but-not-constructed stations. Thus, these should constitute a better control group than other locations.

To illustrate how I define the treatment and control groups, Figure 4 depicts the map of the constructed and planned stations, in addition to buffers around each station. In practice, I define the treatment group as all locations that are within 1.5 kilometers of any Line 1 or BRT station. Then, I define the control group as all locations that are within 3.2 kilometers from any Line 2 or Line 4 station. To avoid contamination from the treated group, I add the restriction that all locations in the control group should also be at least 2 kilometers away from all Line 1 and BRT stations. I use a bigger radius to define the control group because some locations are lost due to contamination which hurts statistical power. However, results are robust to alternative definitions.

[Figure 4 here]

3.3.2 Empirical strategy

I estimate time-event studies by comparing outcomes in locations close to the Metro and BRT stations to outcomes in locations close to yet-to-be-constructed stations. I utilize two data sources: the 1993, 2007 and 2017 Population Censuses, and the 2007-2017 waves of the National Household Survey.

The 1993, 2007 and 2017 Population Censuses. First, due to data limitations I start by performing a time event study with a subset of the data.[22] Then, after showing that pre-trends are parallel, I move to a differences-in-differences specification employing only the 2007 and 2017 data, which allows me to fully leverage the cross-sectional variation of the data.

I start by running the following time event study:

$$\log y_{lt} \quad = \quad \sum_k \beta_k \left[Treat_l \cdot Year_t^k \right] + FE_l + FE_t + \epsilon_{it}$$

where y_{lt} is an outcome constructed from Census data, $Treat_l$ is a treatment dummy taking the value of one if the location l is closest to either the Metro or the BRT station. I interact these variables with year dummies and so β_{2017} measures the effect of the new transit infrastructure on locations that are

[21] I leave out Line 3 because it will run over the same locations as the BRT.

[22] To perform the time-event study I need a panel of locations identified in 1993, 2007, and 2017. The problem is that the match between the 1993 data and the 2007 data is imperfect, partially, because many locations that existed in 2007 did not exist in 1993. Out of 605 locations in either the treatment or control group, I can correctly match 125. Matched locations tend to be at a lower altitude, and are closer to the CBD than unmatched locations, as expected. However, there are no statistically significant differences in the size of the block, in the slope, nor in the overall treatment status. I argue that while this subset of the data is non-random and might introduce a selection effect, it is still reassuring to find parallel pre-trends. Also, out of the 605 locations, I can match 600 in years 2007 and 2017.

closer to the stations, relative to locations that are closer to yet-to-be-constructed stations. $\beta_{1993} = 0$ would be indicative of parallel pre-trends. Finally, I introduce a set of location fixed effects FE_l, in addition to year fixed effects FE_t.

Next, I use the whole sample of locations in 2007 and 2017 to fully leverage the cross-sectional variation and estimate a differences-in-differences model. First, I break the treatment dummy into two: $Treat_l^{Metro}$, which takes the value of one if location l's closest station is a Metro station, and $Treat_l^{BRT}$ if the closest station is a BRT station.[23] Second, I argue that being closer to a station does not guarantee greater access to relevant destinations. Gains in access should be higher in the fringe of the city since these areas were more isolated at the beginning of the sample. Conversely, gains in access in locations closer to the CBD may be less pronounced since workers were already close to valuable destinations. Also, impacts on amenities and other outcomes could potentially be heterogeneous across locations. Thus, I allow for the treatment effects to vary by the distance of households to the CBD. I construct five dummies indicating if the household is located within 0-5, 5-7.5, 7.5-10, 10-12.5 kilometers, and more than 12.5 kilometers.

The 2007-2017 waves of the National Household Survey. I use 2007-2017 rounds of the Peruvian National Household Survey, georeferenced at the block level, to precisely locate households and classify them into treatment and control groups. Then, I run the following time-event study at the household level i and block level l at year t:

$$\begin{aligned}\log y_{il(i)t} &= \sum_k \beta_k \left[Treat_{l(i)}^{Metro} \cdot Time_t^k\right] + \sum_k \beta_k \left[Treat_{l(i)}^{BRT} \cdot Time_t^k\right] \\ &+ \gamma' X_{il(i)t} + FE_{l(i)} + FE_t + \epsilon_{it}\end{aligned} \tag{3}$$

where y_{it} is a household-level outcome such as rents and income.[24] $Treat_{l(i)}^{Metro}$ and $Treat_{l(i)}^{BRT}$ are treatment dummies taking the value of one if the location l is closest to either the Metro or the BRT station. $Time_t$ are a set of time dummies indicating if the data comes from (i) any year preceding 2009 i.e. 2007 or 2008, (ii) the year just before the introduction of the new transit infrastructure i.e 2009, or (iii) any year after the introduction of the new transit infrastructure, i.e. 2010 or after. I set the base category to be the year 2009 since the Metro and the BRT started operating in 2010. X_{1it} is a set of household-level controls comprising age, years of education, civil status, mother's language, and moving history. Finally, X_{2lt} is a set of block-level controls, i.e. the physical size of the block (i.e. the area), the elevation of the block, the median slope, a continuous measure of distance to the central business district, and a continuous measure of distance to the closest station (either constructed or planned) each interacted with year dummies FE_t.

The impact of being close to a station can hide substantial heterogeneity. In fact, gains in access are not uniformly distributed across stations. I allow for the treatment effects to vary by the distance

[23] To control for differential trends among a set of covariates, I include a set X_{lt} of block-level controls, i.e. the physical size of the block (i.e. the area), the elevation of the block, the median slope, a continuous measure of distance to the central business district, and a continuous measure of distance to the closest station (either constructed or planned) each interacted with year dummies FE_t.

[24] Unfortunately, the size of the sample does not allow me to have enough statistical power to explore the heterogeneity in the impact across different types of households.

of households to the CBD. I construct five dummies indicating if the household is located within 0-5, 5-7.5, 7.5-10, 10-12.5 kilometers, and more than 12.5 kilometers. As we will see, pre-trends do not seem to be an issue. Hence, I simplify equation 3 by replacing the time dummies with a dummy variable taking the value of one if the data comes from any year after 2009 and zero otherwise.

3.3.3 Results

Fact 3: Rents and household expenditures tended to increase in the fringe, especially due to the BRT. Table 2 shows the results of estimating equation 3 using survey data. Column (1) depicts the results on a variable constructed by the Bureau of Statistics which considers paid rents plus hypothetical rents. I show that on average there were no statistically detectable pre-trends, but also no average impacts of either the Metro or the BRT.[25]

Moving to measures of household income and expenditure (columns 2 and 3 of Table 2), I find no detectable pre-trends of either the Metro or the BRT. Furthermore, I find that, on average, locations closer to the Metro increased their expenditures by 10.6%. I find a similar coefficient when looking at the effect on income, but it is not statistically different from zero. I do not find any average impact of the BRT. In column 4 of Table 2, I complement previous outcomes with dummies indicating whether household i is poor.[26] Once more, I do not find any pre-trends. Furthermore, I find that the Metro decreased poverty in nearby locations by 4.96 percentage points.

Next, I explore the heterogeneity of the effect across locations. Regarding rents, Panel A of Figure 5 reveals that the estimated treatment effects are greater the farther away you move from the CBD. This is true for both the Metro and especially for the BRT.[27] Panel B of Figure 5 shows the heterogenous impact on expenditures. Regarding the Metro, point estimates tend to be quite homogeneous across locations. For the BRT, I find that expenditures increased especially in locations farther away from the CBD.[28]

[Table 2 here]

[Figure 5 here]

Fact 4: On average, people moved out from the stations. This effect is reinforced in the fringe for the BRT, and in the CBD for the Metro. The basic intuition is that location choices can summarize to some extent the impact of the new infrastructure on households. As it will become clear in the theoretical section of the paper, new transit infrastructure can impact households through many channels. The main channel is the improvement in commuting times leading to greater access to jobs. However, the metro can also affect local amenities, for instance by bringing crowds, crime,

[25] I obtain similar estimates if I use only paid rents instead.

[26] A household is defined as poor if it is unable to consume 2318 kilocalories per day as well as spend on basic services such as clothing, rent, health, transportation, and education, among others (INEI, 2000).

[27] See Panel A of Figure B.4 for impacts on actual paid rents.

[28] See Figure B.4 for impacts on income and poverty rates.

and noise, and thus can also affect the price of the land. All of these will impact the location chosen by households. If we observe population decreasing in treated locations relative to control locations, it would be indicative that the new infrastructure is harming welfare in some way in these locations (e.g. by worsening amenities).

Results for the time-event study utilizing Census data are reported in Table 3. Column (1) of Panel A shows the impact on overall population living in each block. We can see that relative to locations near yet-to-be-constructed stations, population in locations closer to stations shrunk by 12%. Importantly, I do not find evidence for differential pre-trends. In Table B.2, I show the corresponding differences-in-difference specification using the whole sample of 2007 and 2017 locations. Results are similar albeit coefficients are somewhat smaller. While the negative impact may seem surprising given the improvement in commuting times, it is not if we consider that rents were increasing especially in the fringe, and that amenities could be deteriorating.

Next, Figure 6 shows the heterogeneity of the impact across locations. On the one hand, estimates reveal that the Metro decreased the population located in the CBD. On the other, the BRT had a negative impact on locations in the fringe. This could be because the BRT and the Metro affected access to opportunities, rents, and amenities in different magnitudes across locations. For instance, given that rents increased mainly in the fringe but not in the CBD, and that the Metro had a negative impact on population only in the CBD, this would imply that the Metro deteriorated amenities in the CBD.

[Table 3 here]

[Figure 6 here]

Fact 5: Males are better off in intermediate locations (between the CBD and the fringe) but married women are better off in the fringe. In Panel A of Table 3, I exhibit the impact on other outcomes. Column (2) shows the impact on the ratio of singles to married, while column (3) shows the impact on the ratio of female to males. Estimates indicate that singles were more benefited than married households: in locations closer to stations relative to locations near yet-to-be-constructed stations, the population of singles relative to married increased by 8.1%. However, on average, the ratio of males and female stayed constant. Finally, column (4) shows the impact on the number of dual earner households relative to male breadwinner households. On average I do not find any impact. Also note that, importantly, pre-trends are always statistically indistinguishable from zero.[29]

Panel B of Figure 6 reveals that locations around 7.5-10 km from the CBD experienced a decrease of singles relative to married for both the Metro and the BRT. However, both at the CBD and at the fringe, some increases were detectable for the BRT. Panel A of Figure 7 shows the impact on the ratio of females to males. I find that women were more prone to leave locations around 7.5-10 km from the CBD. This implies that males were better off in intermediate locations. Finally, in Panel B of Figure 7 the ratio of dual-earners to male breadwinner households increased in the fringe. So,

[29]In Table B.2, I show the corresponding differences-in-difference specification using the whole sample of 2007 and 2017 locations.

while single households are equally or better off in the fringe (and worse off between 7.5-10 km from the CBD), dual-earner households are better off compared to male breadwinner households, implying that opportunities to women—and in particular to married women—expanded in the fringe.

[Figure 7 here]

In conclusion, first, these results reveal that the new transit infrastructure fostered households to move out from areas closer to the stations. However, this impact was heterogeneous across household groups. For instance, results suggest that males' welfare expanded in intermediate locations, whereas females' welfare did so in the fringe, especially married women's welfare. Second, as it will become clear with the model, this rich heterogeneity comes from a combination of facts related to the geography of improvements in commuting times, the initial distribution of households and jobs across space, potential changes in amenities, and other general equilibrium impacts such as on rents and on wages. The model allows me to tease out the impacts that come exclusively from having better commute times (and the resulting GE impacts). Third, comparing stations to yet-to-be-constructed stations seems to take care of some endogeneity concerns since I do not find strong evidence for pre-trends. Not a single pre-trend coefficient is significant. Fourth, these results point towards the need of a model that takes the geography of Lima very seriously since impacts vary across locations.[30]

Finally, another word of caution when interpreting these impacts is that the introduction of the new transit infrastructure may have induced households to relocate to other parts of the city, as the Census data seems to suggest. In other words, the impacts I estimate are conflating partial and general equilibrium effects. It is hard to extract definitive conclusions from this analysis since the introduction of these transit investments have potentially affected the whole city, blurring the lines between what control and treatment groups are. To better understand these issues and tease out the different forces that are interacting in this context, in the following sections I start building intuition by developing a partial equilibrium model of commuting in married households. Then, in the quantitative section of the paper I set up a general equilibrium model of city structure based on Ahlfeldt, et al. (2015) and Tsivanidis (2021) but with the addition of interdependent commuting choices within couples.

4 Theory

To build intuition, I set up a simple model of commuting among married households. Then, I study the model's implications in a simple environment. In Section 5, I embed this choice structure into a spatial model of city structure.

[30] One could argue that Lima cannot be fully described by a monocentric model were only one CBD exists. Hence, the geographical dimension in the model should be flexible enough to account for this heterogeneity.

4.1 Microfoundation of a household utility for married households

Let V_{ij}^{k} be the utility of spouse k living in i and working in j, such that:

$$\begin{aligned} V_{ij}^{k} &= \log \frac{C_{k,i}^{\beta^k} H_{R_i}^{1-\beta^k}}{d_{ij}^{k}} \\ \text{s.t.} \quad & PC_{k,i} + PC_{k',i} + r_{R_i} H_{R_i} = w_{k,j} + w_{k',j'} \end{aligned}$$

where $C_{k,i}$ is the consumption of family member k of the final good, H_{R_i} is the household consumption of residential floorspace. Note that residential floorspace is a public good within the couple. d_{ij}^{k} is an iceberg commuting cost $d_{ij}^{k} = \exp(\kappa_k t_{ij})$. t_{ij} is the time it takes to commute to j from i, and κ_k is the rate at which commute times are transformed into commute costs. P is the price of the final good, which does not depend on the origin location, because in the quantitative exercise I assume that the final good is freely traded within the city. r_{R_i} is the rental value of the residential floorspace. Finally, $w_{k,j}$ is the wage received by member k when working at j, whereas $w_{k',j'}$ is the wage received by the other member at j'.

Following the literature on family economics, I assume households only choose efficient outcomes (Browning et al., 2011). That is, for any given vector of prices and wages $(P, r_{R_i}, w_{k,j}, w_{k',j'})$ and location characteristics $(d_{ij}^{k}, d_{ij'}^{k'})$, an allocation $(C_{k,i}, H_{R_i})$ is efficient if there exists a feasible $\bar{V_{ij}}^{k'}\left(P, r_{R_i}, w_{k,j}, w_{k',j'}\right)$ such that $(C_{k,i}, H_{R_i})$ solves the problem:

$$\begin{aligned} \max_{C_{k,i}, H_{R_i}} \quad & V_{ij'}^{k} \\ \text{s.t.} \quad & PC_{k,i} + PC_{k',i} + r_{R_i} H_{R_i} = w_{k,j} + w_{k',j'} \\ \text{and} \quad & V_{ij}^{k'} \geq \bar{V_{ij}}^{k'}\left(P_i, r_{R_i}, w_{k,j}, w_{k',j'}\right) \end{aligned}$$

This maximization problem can be broken into two stages. In the first stage, households choose the optimal level of residential floorspace H_{R_i} and the disposable income allocated to each spouse, $x_{k,i}, x_{k',i}$. In the second, each spouse allocates their income into the consumption of the final good. Then, the first stage becomes:

$$\begin{aligned} \max_{x_{k,i}, H_{R_i}} \quad & W_{ijj'} \equiv \lambda V_{ij'}^{k'} + V_{ij}^{k} \\ \text{s.t.} \quad & x_{k,i} + x_{k',i} + r_{R_i} H_{R_i} = w_{k,j} + w_{k',j'} \equiv y_{ijj'} \end{aligned}$$

The first order conditions yield two sets of equations. The first set equalizes the marginal utility derived from the consumption of one spouse to that of the other spouse: $\lambda \frac{\partial V_{ij}^{k'}}{\partial x_{k',i}} = \frac{\partial V_{ij'}^{w}}{\partial x_{k,i}}$. The second set, also known as the Bowe-Lindahl-Samuelson conditions, says that members' marginal willingness to pay should be equalized to the marginal price of residential floorspace: $\frac{\partial V_{ij}^{k}/\partial H_{R_i}}{\partial V_{ij}^{k}/\partial x_{k,i}} + \frac{\partial V_{ij'}^{k'}/\partial H_{R_i}}{\partial V_{ij'}^{k'}/\partial x_{k',i}} = r_{R_i}$.

In the quantitative exercise, I abstract away from bargaining considerations within households and only consider the feature that spouses share the same budget. For that reason, I assume that λ, typically thought as a measure of bargaining power, is exogenous and equal to one so that the household weights spouses' utility equally.[31] Moreover, I assume that spouses have the same preferences over consumption and housing and so $\beta^k = \beta^{k'} = \beta$. These assumptions and the first order conditions yield the following solutions to the maximization problem: $x^k_{ijj'} = x^{k'}_{ijj'} = \frac{\beta}{2} y_{ijj'}$ and $r_{R_i} H_{R_i} = (1-\beta) y_{ijj'}$. Plugging these back into the household utility function, I get that:[32]

$$\tilde{W}_{ijj'} \propto \left(d^k_{ij} d^{k'}_{ij'} \right)^{-1} \left(\frac{w_{k,j} + w_{k',j'}}{P^{\beta} r_{R_i}^{(1-\beta)}} \right) \tag{4}$$

This household utility postulates that the household dislikes the commuting performed by both members, which is a natural outcome from the fact that the household cares about the average utility of members. Moreover, greater income (or wages) and lower prices provide greater household utility.

4.2 Implications

Marginal rates of substitution. In this section, I explain what are the implications of this choice structure. Specifically, I ask two questions.[33] First, say that the times from origin i to destination j' improve for the wife by one minute. How much can the husband's commuting times increase while keeping the utility constant? Second, keeping husband's commuting times constant, what is the wage of the husband that the household would be willing to give up in order to improve the commuting time of the wife by one minute?

To answer the first question, we can take a total derivative of the utility function with respect to commute times of spouses and equalize it to zero. This yields $dt^h_{ij} = -\frac{\kappa_h}{\kappa_w} dt^w_{ij}$ where h and w index the husband and the wife. This result means that the household can substitute the time freed up from the wife's commuting with $\frac{\kappa_h}{\kappa_w} dt^w_{ij'}$ minutes in husband's commuting. $\frac{\kappa_h}{\kappa_w}$ acts as an exchange rate at which the time of the wife can be substituted with the time of the husband. This structure of the commute costs will deliver gravity equations for commute flows, greatly simplifying the process of solving and estimating the model.

To answer the second question, I follow a similar strategy. To decrease the commute times of the wife by one minute, the household would be willing to give up income $w_{h,j} + w_{w,j'}$ by κ_w log points: $d\log(w_{h,j} + w_{w,j'}) = -\kappa_w dt^w_{ij'}$. This is a feature also present in the class of spatial models with iceberg costs and constant returns to scale utility functions. The difference in this setting is that I generalize it: income depends on the wage of the two spouses, and so each spouse's commuting is responsible of sourcing a share of the household's income. Noting that $d\log(w_{h,j} + w_{w,j'}) = \frac{1}{w_{h,j} + w_{w,j'}} dw_{h,j}$, then $\frac{dw_{h,j}}{w_{h,j}} = \kappa_w \frac{w_{h,j} + w_{w,j'}}{w_{h,j}} dt^w_{ij'}$. Thus, as $w_{w,j'}$ increases, households are willing to give up more of the husband's wage in relative terms to decrease commute costs, since the household has already secured the income of the wife.

[31] Under this assumption, households behave as a single entity that makes all choices, i.e. households behave as unitary households.

[32] I have taken the exponential function to both sides, taken the root square and redefined the meaning of κ^k. These preferences should represent the same preferences ordering since all operations were monotonic transformations.

[33] For ease of explanation and because this is the way I take the model to the data, I assume that married households are comprised of a woman (the wife) and a man (the husband).

Gravity equations. Assuming that individuals in household ω have heterogeneous preferences over their commute destinations, $\epsilon_{k,j}(\omega)$, and $\epsilon_{k',j'}(\omega)$. Solving the maximization problem above yields:

$$W_{ijj'}(\omega) \propto \frac{\epsilon_{k,j}(\omega)\epsilon_{k',j'}(\omega)}{d_{ij}^{k}d_{ij'}^{k'}}\left(\frac{w_{k,j}+w_{k',j'}}{P^{\beta}r_{R_i}^{(1-\beta)}}\right) \tag{5}$$

I assume preference shocks are Frechet distributed with shape parameters θ_k and $\theta_{k'}$. Thus, conditional on spouse k''s destination j', the household evaluates the utility it would get from sending k to each possible destination and chooses the one that maximizes it. Similarly, conditional on spouse k's destination j, the household evaluates all possible destinations of member k'. By standard Frechet operations, I get that:

$$\pi_{j|ij'}^{k} = \frac{W_{ijj'}^{\theta_k}}{\sum_l W_{ilj'}^{\theta_k}} \quad \text{and} \quad \pi_{j'|ij}^{k'} = \frac{W_{ijj'}^{\theta_{k'}}}{\sum_l W_{ijl}^{\theta_{k'}}}$$

For simplicity in this exposition, I assume that $\theta_k = \theta_{k'} = \theta$. In the quantitative exercise, I allow for the shape parameters to vary across household members. Under this assumption, unconditional commuting probabilities become:

$$\pi_{j|i}^{k} = \frac{\sum_l W_{ijl}^{\theta}}{\sum_{l'}\sum_l W_{ill'}^{\theta}} \quad \text{and} \quad \pi_{j'|i}^{k'} = \frac{\sum_l W_{ilj}^{\theta}}{\sum_{l'}\sum_l W_{ill'}^{\theta}}$$

Note that both sets of commuting probabilities have a gravity structure: there is a linear relationship between the log of commuting probabilities of member k and his or her commuting times. The product $\theta\kappa_k$ becomes the commuting elasticity. This result allows me to solve and estimate the model more easily.

Comparative statics. I compute the impact of improving the commuting times of spouse k' on the commuting probabilities of spouse k. First, note that improving k''s commuting times does not affect k's conditional commuting probabilities. This is because when k''s commuting times to j' improve, the utility received in every destination of k changes proportionally to the initial utility: $\frac{dW_{ijj'}}{dt_{ij'}^{k'}} = -\kappa_{k'}W_{ijj'}, \forall j$.

However, k's unconditional probabilities are affected as the relative ranking of utility $W_{ijj'}$ across destinations j and j' is reordered thanks to the improvement in commuting times: $\pi_{j|i}^{k} = \frac{\sum_l (W_{ijl}+1[l=j']\kappa_{k'}W_{ijj'})^{\theta}}{\sum_{l'}\sum_l (W_{ill'}+1[l'=j']\kappa_{k'}W_{ilj'})^{\theta}}$. This is as if households, by having more time at their disposal, reallocate spouses across destination such that utility is maximized. Commuting to location j will increase if the resulting numerator $\sum_l (W_{ijl}+1[l=j']\kappa_{k'}W_{ijj'})^{\theta}$ is greater than the numerator of the remaining options. Intuitively, this will happen when the location j features high wages or low commute costs, as the impact on $W_{ijj'}$ is proportional. If wages in j are too low or commute costs to j too high, utility maximization leads households to decrease the probability at which they send spouses k to such destination.

Mathematically, how do k's unconditional probabilities change when k''s times improve? Taking a derivative with respect to $t_{ij'}^{k'}$, I get:

$$\frac{\partial \pi_{j|i}^{k}}{\partial t_{ij'}^{k'}} = -\theta\kappa_{k'}\pi_{j|i}^{k'}\left(\pi_{j|ij'}^{k} - \pi_{j|i}^{k}\right) = -\theta\kappa_{k'}\pi_{j|i}^{k}\left(\pi_{j'|ij}^{k'} - \pi_{j'|i}^{k'}\right) \tag{6}$$

Spouses k will increase their probability of commuting to location j if the probability that spouses k use the route j given that their spouses k' use the improved route j' is higher than the unconditional probability that spouses k use the route j. In other words, the probability that k commutes to j will increase when initial conditional commuting probabilities $\pi_{j|ij'}^{k} = (\frac{w_{k,j}+w_{k',j'}}{d_{ij}^{k}})^{\theta}/\sum_l(\frac{w_{k,l}+w_{k',j'}}{d_{il}^{k}})^{\theta}$ are high enough. The second equality follows from an application of the Bayes theorem.

How improvements in commuting time affect income? Defining spouses k's income at origin i as the sum of wages across destinations $y_i^k = \sum_l w_{k,l}\pi_{k,l|i}N_i^k$, where N_i^k is the number of spouses k living in origin i. Differentiating it with respect to commute times between origin i and a destination j' of both spouses k and k', I get:

$$-\frac{\partial}{\partial t_{ij'}}\log y_i^k = \underbrace{\theta\kappa_k\frac{w_{k,j'}\pi_{j'|i}^{k}}{\sum_l w_{k,l}\pi_{l|i}^{k}}\left(1-\frac{\bar{y}_i^k}{w_{k,j'}}\right)}_{\text{direct use}} + \underbrace{\theta\kappa_{k'}\sum_l\frac{w_{k,l}\pi_{l|i}^{k}}{\sum_l w_{k,l}\pi_{l|i}^{k}}\left(\pi_{j'|il}^{k'} - \pi_{j'|i}^{k'}\right)}_{\text{indirect use}} \tag{7}$$

The effect on total income in origin i for spouse k depends on two terms: the direct use channel, and the indirect use channel (or within household spillovers). The direct use channel will be positive if $w_{k,j'} > \bar{y}_i^k$, where $\bar{y}_i^k = \sum_j\left(w_{k,j}\pi_{j|i}^{k}\right)$. In other words, when times between origin i and destination j' improve, aggregate income in origin i increases if wages at the destination that is now more accessible are higher than the average income in origin i. The intuition is that income should increase only if the new destination j' is a high-wage location. Otherwise, if the new destination is a low-wage location, improvements in commuting times may induce some workers to give up some of their earnings to have a shorter commute.[34] How much would it increase depends on three terms. First, it depends on the share of income that destination j' accounts for, $\frac{w_{k,j'}\pi_{j'|i}^{k}}{\sum_j\left(w_{k,j}\pi_{j|i}^{k}\right)}$. Second, it depends on the difference between wages $w_{k,j'}$ and average income $\bar{y}_i^k$. Finally, it depends on the commuting elasticity $\theta\kappa^k$.

The indirect use channel indicates that when commute costs for spouses k' decrease, then this affects the households' allocation of spouses k across destinations as explained above. When this happens, the origin i stops receiving the income from locations where the spouse k is no longer working (i.e. when $\pi_{j'|il}^{k'} < \pi_{j'|i}^{k'}$), but starts receiving the income from locations where spouses k are commuting to (when $\pi_{j'|il}^{k'} > \pi_{j'|i}^{k'}$). We need to sum across all locations l to account for all spillovers generated by spouses k' to spouses k working in locations l. Finally, note that if choices are independent, as in the case of standard spatial models, then conditional commuting probabilities become equal to unconditional probabilities, collapsing the indirect channel to zero.

[34] Note, however, that new metro lines improve commuting times between a set of origin and destination locations. The derivative I have taken only considers an improvement between a particular origin and a particular destination, assuming all other times remain the same which is physically impossible. If if its quicker to reach j' from i, it will also be quicker to reach a neighboring location j''.

In sum, in this section I have shown the microfoundation of a utility function tailored for married households. Then, I have explained the major implications of this choice structure in terms of commuting and have shown how improvements in commute times may impact income across origins. The main idea is that there are two channels through which improved commute times can affect households. One is the direct use channel, and the other one is the indirect use channel (or within household spillovers). In the next section, I summarize how I bring this model to the data.

5 Bridging Theory and Data

This section starts by summarizing the main components of the quantitative model. Afterwards, it describes the calibration and estimation procedures. Finally, it shows evidence about the model's performance when matching untargeted moments.

5.1 Quantitative Model of City Structure and Couples

I embed the choice structure for married households into a spatial model of city structure. Not only I incorporate married households, but also single households. Moreover, I consider other margins studied in the literature through which new transit infrastructure can affect households to successfully unite the theory with the data.

5.1.1 Environment

This is a model in which the internal structure of the city is driven by a tension between agglomeration forces (production and positive residential externalities) and dispersion forces (commuting costs, negative residential externalities and an inelastic supply of land). This tension is heterogeneous across household groups, that is, among single males and females, and especially among married couples.

The city consists of a set of discrete locations indexed by $i = 1, ..., S$. These locations differ in their commute times to other locations, their housing floorspace, households productivity, amenities, and industries' productivity as I explain later.

Moreover, the city is populated by a fixed measure of households N_k, where $k \in \{m, f, mf\}$. If $k = m$ or $k = f$ the household is composed of a single male worker or a single female worker, respectively. If $k = mf$ the household is a couple.[35] For simplicity, singles and husbands always work in this model.[36] Thus, there are five groups of workers: single males, single females, married males in male-breadwinner households, and males and females in dual-earner households.

Furthermore, each location houses multiple industries which produce a consumption good using labor and commercial floorspace. Each location-by-industry's consumption good is aggregated into a final good with CES preferences. Males and females are imperfect substitutes in the production

[35] To simplify, in this paper couples are comprised of a male and a female.

[36] A simplifying assumption usually done in the literature on family economics is that the husband always works. Moreover, in the data female breadwinner households constitute just 1.3% of all households. This is why I turn off this margin: it does not seem to be quantitatively important but would add greater complexity to the model.

function. Finally, landowners choose how to allocate a fixed amount of floorspace across residential and commercial use.

Single households

A. Preferences in single households. Singles derive Cobb-Douglas utility from the consumption of a freely traded numeraire good ($C^g(\omega)$) with parameter β^g, from the consumption of residential floorspace $H_{R^g}(\omega)$ with parameter $1-\beta^g$, and from an amenity accounting for the average preference of each group to live in location i (u_i^g). Moreover, individuals experience disutility from commuting in an iceberg fashion, $d_{ij}^g \geq 1$, where $d_{ij}^g = \exp(\kappa^g t_{ij})$, and t_{ij} is the time it takes to commute between locations i and j. The parameter κ^g governs the size of these commuting costs for each gender. Finally, workers are heterogeneous in their preferences for working in location j, $\epsilon_j^g(\omega)$, and for living in residence i, $\nu_i^g(\omega)$. Hence, singles need to decide where to work based on the trade-off between the possibility of higher wages but longer commutes. Similarly, they need to decide where to live by balancing out the expected income against housing costs and amenities. Concretely, indirect utility is given by:

$$V_{ij}^g(\omega) \quad \propto \quad u_i^g \nu_i^g(\omega) P^{-\beta^g} r_{R_i}^{\beta^g-1} \left[\frac{w_{jg}\epsilon_j^g(\omega)}{d_{ij}^g} \right] \tag{8}$$

where P is the price of the freely traded final good, and r_{R_i} is the rental value of residential floorspace. Solving backwards, conditional on the locations that singles live and work, they choose their consumption and housing allocations. Afterwards, conditional on their residence location, they choose their workplace location. Finally, they decide where to live. Panel A of Figure B.5 depicts the timing assumptions.

B. Commuting. Conditional on where to live, singles (indexed by ω) draw a vector of location-specific preferences for work across the city, $\epsilon_j^g(\omega)$, iid from a Frechet distribution $F(\epsilon_j) = \exp\left(-T^g \epsilon_j^{-\theta^g}\right)$. From standard Frechet operations, we know that:

$$\pi_{j|i}^g \quad = \quad \frac{\left(\frac{w_{jg}}{d_{ij}^g}\right)^{\theta^g}}{\Psi_{R_i^g}} \tag{9}$$

where the probability singles g choose destination j given origin i is: $\pi_{j|i}^g$. Moreover, $\Psi_{R_i^g} = \sum_{l=1}^S \left(\frac{w_{lg}}{d_{ij}^g}\right)^{\theta^g}$ is a market access term indicating how close are jobs to people. Thus, one key trade-off is that individuals are attracted to locations paying a high wage but dislike commuting long distances. That is, they compare the net wage they would get from commuting to particular destination with other available options embodied in $\Psi_{R_i^g}$. The dispersion of preferences is specific for each gender, and it determines how sensitive singles are to changes in commute costs. For example, when the dispersion of preferences shocks is low (i.e. high θ^g), choices are more sensitive to commute costs.

C. Residence Location. Before singles draw a vector of location-specific preferences, they draw a value for residence preferences, $\nu_i^g(\omega)$, and decide where to live. $\nu_i^g(\omega)$ is distributed Frechet with shape parameter $\eta^g > 1$ and average E^g. So, singles choose to live in i only if they attain a higher expected utility than in all other locations. Then, the probability that single ω lives in location i is:

$$\pi_i^g = \frac{\left(\frac{r_{R_i}^{1-\beta^g}}{u_i^g \Psi_{R_i^g}^{1/\theta^g}}\right)^{-\eta^g}}{\sum_l \left(\frac{r_{R_l}^{1-\beta^g}}{u_l^g \Psi_{R_l^g}^{1/\theta^g}}\right)^{-\eta^g}} \tag{10}$$

This expression says that, when deciding where to live, singles will balance out locations with better amenities and greater access to jobs against the cost of housing. This is another key trade-off in the model.

Married households

A. Preferences in married households. Individuals in married households derive utility from the same components as singles.[37] Thus, preferences for married households can be microfounded following a similar procedure as in Section 4.1.

I additionally incorporate the labor participation choice of women, which might be a relevant margin in this context. To do so, I assume that households can engage in household production with location-specific productivity $\xi_{i\ell}^{mf}$, which depends on whether both spouses work ($\ell = 1$) or just one ($\ell = 0$). There could also be ideosyncratic reasons why a household prefers to have both spouses working in the labor market, which are embodied in α_ℓ^{mf}. I further assume that males and females have the same preference over residential location ν_i^{mf} and derive the same utility from local amenities u_i^{mf}. They only differ on their preferences over destinations for work, ϵ_j^k, and commuting time d_{ij}^k. Thus, indirect utility of a couple ω living in i, with the husband working at j and the wife in j' (if she works) is given by:

$$W_{i\ell jj'}^{mf}(\omega) \quad \propto \quad u_i^{mf} \nu_i^{mf}(\omega)\, \xi_{i\ell}^{mf} \alpha_\ell^{mf}(\omega)\, \frac{\epsilon_j^h(\omega)}{d_{ij}^h} \left(\frac{\epsilon_{j'}^w(\omega)}{d_{ij'}^w}\right)^{\ell} P^{-\beta^{mf}} r_{R_i}^{\beta^{mf}-1} y_{\ell jj'}^{mf}(\omega)$$

where household ω's income is $y_{\ell jj'}^{mf}(\omega) = w_{jh} + w_{j'w}\ell(\omega)$. Notice that a couple ω performs the same choices as singles. Additionally, they have to choose whether the wife is going to work ($\ell = 1$ if she works, $\ell = 0$ if she stays at home), and the location where she is going to work (indexed by j'). Thus, married couples face additional trade-offs. First, if the wife decides to work, households will dispose of more income to spend in housing and the final good. However, if the wife stays, households will use her unit of labor to produce a household good with location-specific productivity $\xi_{i\ell=0}^{mf}$. Second, there is a potential trade-off between commute time and income, as explained in the previous section.

Solving backwards, married households start by deciding how much they consume in the final good

[37] I indicate if I am referring to husbands using the subscript h and to wives using the subscript w.

and in residence floorspace, conditional on the workplace location of both spouses and the residential location. Afterwards, conditional on the wife's labor status and the wife's workplace location, households decide the workplace of the husband. Then, they decide the workplace of the wife. If the wife does not participate, households only decide the workplace location of the husband. Next, conditional on households residence location, households compare the utility they would receive if the wife works against the utility they would receive if the wife stays at home performing household production. Based on this comparison, households decide the labor status of the wife. Finally, households choose their residence location. Panel B of Figure B.5 depicts the timing assumptions.

I start with the case where the wife participates in the labor market. Afterwards, I solve for the case where the wife stays at home performing household production. Then, I determine labor participation and residence choices.

B. Husbands' Commuting Choice in Dual-Earner Households. I assume for simplicity that households choose the husband's workplace after knowing the workplace of the wife. Households compare the utility they would receive if the husband works in a particular location, against the utility they would derive if he works in some other location, all conditional on where the wife works. Hence, we are interested in computing the *conditional* probability that married men living i commute to j given that their wives work and commute to j', i.e. $\pi^h_{j|i,\ell=1,j'} = P\left[W_{i,\ell=1,j,j'}(\omega) \geq \max_{l\neq j} W_{i,\ell=1,l,j'}(\omega)\right]$, where $\ell = 1$ indicates that the wife participates in the labor market. Simple manipulations result in the following expression:

$$\pi^h_{j|i,\ell=1,j'} = \frac{\left(\frac{w_{jh}+w_{j'w}}{d^h_{ij}}\right)^{\theta^h}}{\Phi^h_{R_{i,\ell=1,j'}}} \tag{11}$$

where $\Phi^h_{R_{i,\ell=1,j'}} = \sum_l \left(\frac{w_{lh}+w_{j'w}}{d^h_{il}}\right)^{\theta^h}$ is a market access term for husbands. It indicates how close are jobs to husbands but conditional on the wife's wage. It can also be thought as the expected utility of households when the wife works at j', but before households have decided where to send the husband. As in single households, married men are more likely to commute to a location when it pays a high income net of commute cost. However, now income includes the wage earned by the wife when she works at j'. Finally, the sensitivity of employment decisions to commute costs is governed by the dispersion of taste shocks. So, differences in tastes control the extent to which individuals are willing to bear high commuting costs.

C. Wives' Commuting Choice in Dual-Earner Households. To decide where the wife should work, the household compares the expected utility it would receive across all possible options. In particular,

$$\pi^w_{j'|i,\ell=1} = \frac{\left(\frac{\left(\Phi^h_{R_{i,\ell=1,j'}}\right)^{1/\theta^h}}{d^w_{ij'}}\right)^{\theta^w}}{\Phi^w_{R_{i,\ell=1}}} \tag{12}$$

where $\Phi^w_{R_{i,\ell=1}} = \sum_l \left(\frac{\left(\Phi^h_{R_{i,\ell=1,l}}\right)^{1/\theta^h}}{d^w_{il}} \right)^{\theta^w}$ is a market access term for dual-earner households. Dual-earner households can have greater access to jobs either through the husband or through the wife, and so $\Phi^w_{R_{i,\ell=1}}$ acts as an average of the two. This equation says that wives will commute to locations where they expect a higher household income, $\Phi_{R^h_{i,\ell=1,j'}}$, net of their own disutility of commuting, $d^w_{ij'}$.

D. Husbands' Commuting Choice in Male Breadwinner Households. In the case of male breadwinner households, a similar logic as in the case of single households delivers the following expression for commuting probabilities:

$$\pi^h_{j|i,\ell=0} = \frac{\left(\frac{w_{jh}}{d^h_{ij}}\right)^{\theta^h}}{\Phi^h_{R_{i,\ell=0}}} \tag{13}$$

So, in terms of commuting, married males in breadwinner households and single males behave quite similarly, although with potentially different commuting elasticities.

E. Labor Force Participation. Conditional on their residence, households compare the expected utility they would obtain if the wife works or if she stays at home performing household production. Since households are heterogeneous in their preference for a stay-at-home wife, and assuming this heterogeneity is distributed Frechet with shape parameter ν, the probability that a couple decides to be a dual-earner household is:

$$\mu_i = \frac{\left[\left(\Phi^w_{R_{i,\ell=1}}\right)^{1/\theta^w}\right]^{\nu}}{\Psi^{mf}_{R_i}} \tag{14}$$

where $\Psi^{mf}_{R_i} = \left[\left(\Phi^w_{R_{i,\ell=1}}\right)^{1/\theta^w}\right]^{\nu} + \left[\xi_{i\ell=0}\left(\Phi^h_{R_{i,\ell=0}}\right)^{1/\theta^h}\right]^{\nu}$. Household productivity in domestic goods $\xi_{i\ell=1}$ is normalized to one, and hence $\xi^{mf}_{i\ell=0}$ indicates how much more productive are households on domestic production in location i when the wife stays at home relative to the situation in which she participates in the labor market.[38] This expression says that if both spouses have opportunities to work by having access to nearby jobs, then the likelihood of both working increases relative to the situation where only males have greater access to jobs. In this respect, ν measures the inverse of the dispersion in households tastes about wives' labor force participation. When the dispersion of preferences is low (i.e. high ν), choices are more sensitive to small changes in the expected utility of wives staying at home for household production, for instance. Thus, this parameter governs the labor supply elasticity of women's extensive margin in married households.

F. Residence Location. Couples choose to live in i only if they attain a higher expected utility than in

[38] Since $\left(T^h\right)^{1/\theta^h}\Gamma\left(\frac{\theta^h-1}{\theta^h}\right)$ only affects the scale of welfare I have normalized this constant to one hereafter.

all other locations. Then the probability household ω in location i is:

$$\pi_i^{mf} \quad = \quad \frac{\left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_i^{mf}\Psi_{R_i^{mf}}^{1/\nu}}\right)^{-\eta^{mf}}}{\sum_l \left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_l^{mf}\Psi_{R_l^{mf}}^{1/\nu}}\right)^{-\eta^{mf}}} \tag{15}$$

Hence, not only the location elasticity η is different in single households than in married couples, but also the married have other considerations when deciding where to live. That is, they consider locations where on average both spouses have better labor opportunities. If only males have good opportunities then they look for locations with greater household productivity. Thereby, this directly links females and males market conditions, a mechanism absent for singles.

Aggregation

I aggregate the supply of residents and of workers to jobs. Once commuting probabilities are known, one can compute aggregate income by origin by weighing wages in each destination with the corresponding commuting probability.

Supply of residents. Equation 10 characterizes the location choice of singles, whereas equation 15 does it for married couples. Given the aggregate number of households-by-group we can compute the supply of households in each location i:

$$N_{R_i^g} = \pi_i^g N^g, g \in \{m, f\} \quad \text{and} \quad N_{R_i^{mf}} = \pi_i^{mf} N^{mf} = \underbrace{(1-\mu_i)\,\pi_i^{mf} N^{mf}}_{N_{R_{i,0}^{mf}}} + \underbrace{(\mu_i)\,\pi_i^{mf} N^{mf}}_{N_{R_{i,1}^{mf}}} \tag{16}$$

Note that given the number of married couples in location i, $N_{R_i^{mf}}$, I endogenously determine which are dual-earner $N_{R_{i,1}^{mf}}$ or male breadwinner $N_{R_{i,0}^{mf}}$ households.

Supply of workers to jobs. After computing how many households live in block i as given by equation 16, and the probability of commuting to any block as given by equations 9, 11, 12, and 13, we can calculate the supply of workers of gender g to any location:

$$\begin{aligned} L_{F_{jm}} &= \sum_i \left\{\pi_{j|i}^m N_{R_i^m}\right\} + \sum_i \left\{\pi_{j|i,\ell=0}^h N_{R_{i,0}^{mf}}\right\} + \sum_i \left\{\pi_{j|i,\ell=1}^h N_{R_{i,1}^{mf}}\right\} \\ L_{F_{jf}} &= \sum_i \left\{\pi_{j|i}^f N_{R_i^f}\right\} + \sum_i \left\{\pi_{j|i,\ell=1}^w N_{R_{i,1}^{mf}}\right\} \end{aligned} \tag{17}$$

where the commuting probability of males in dual-earner households $\pi_{j|i,\ell=1}^h$ is given by

$$\pi_{j|i,\ell=1}^h \quad = \textstyle\sum_l \pi_{j|i,\ell=1,l}^h \cdot \pi_{l|i,\ell=1}^w \tag{18}$$

Finally, aggregate income in origin i of individuals belonging to group k becomes $y_i^k = \sum_l w_{k,l}\pi_{il}^k N_i^k$. Given Cobb-Douglas preferences, we know how much income is destined into the final good and into housing.[39]

5.1.2 Production

I model the production side using standard assumptions in the literature.

Technology

Following an Armington assumption, there are $s \in 1, ..., K$ industries which produce varieties differentiated by location under perfect competition. These goods are aggregated with CES preferences by consumers with an elasticity of substitution of $\sigma_D > 1$. Following Tsivanidis (2021), firms produce using a Cobb-Douglas technology over labor and commercial floorspace:

$$Y_{js} = A_{js} N_{js}^{\alpha_s} H_{F_{js}}^{1-\alpha_s}$$
$$\text{where} \quad N_{js} = \left(\sum_g \alpha_{sg} \tilde{L}_{F_{jgs}}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

Industries aggregate workers of different genders using a CES aggregator over each gender's labor with an elasticity of substitution of σ.[40] Note as well that the total labor share is the sum of each gender's labor share: $\alpha_s = \sum_g \alpha_{sg}$. Finally, A_{js} is the productivity of location j for firms in industry s. Importantly, industries differ in the intensity in which they use different type of workers α_{sg}. The effect of improving access to workers of different genders hinges on which type of jobs they are gaining access to. Assuming perfect competition, prices become $p_{js} = W_{js}^{\alpha_s} r_{F_j}^{1-\alpha_s}/A_{js}$, where $W_{js} = \left(\sum_g \alpha_{sg}^{\sigma} w_{jg}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ is the cost of labor for industry s in location j.

5.1.3 Housing

I also model the housing market in a straightforward way.

Market clearing

Each location is characterized by a fixed amount of floorspace H_i. A fraction ϑ_i is allocated to residential use and $1 - \vartheta_i$ to commercial use. Market clearing requires the supply of residential floorspace, $\vartheta_i H_i r_{R_i}$, to equal demand, $H_{R_i} = \sum_k H_{R_i^k}$, where k indexes household groups. This is the same for commercial floorspace. The supply of commercial floorspace, $(1 - \vartheta_j) H_j r_{F_j}$, must equal the demand from the production-side, $H_{F_j} = \sum_s H_{F_{js}}$.

[39]To avoid the inefficiency generated by absentee landlords, I am assuming that landlords spend all their income in consuming the final good.

[40]Hence, in equilibrium, a gender gap between males and females wages may arise in each location. Also, I assume that single and married males are perfect substitutes. Single and married females are perfect substitutes too. Hence, conditional on a particular destination, married and single individuals of a particular gender will receive the same wage. However, income may vary across these groups depending on their preferred origins and destinations.

Floorspace use allocation.

As in Ahlfeldt et al. (2015) landowners allocate floorspace to its most profitable use. When convenient they allocate ϑ_i to residential. They receive r_{R_i} per unit. For commercial use, land regulations limit the return to each unit, hence they receive $(1-\tau_i)r_{F_i}$. Then:

$$\begin{aligned} \vartheta_i = 1, &\quad \text{if } r_{R_i} > (1-\tau_i)r_{F_i} \\ (1-\tau_i)r_{F_i} = r_{R_i}, &\quad \forall \{i : \vartheta_i \in (0,1)\} \\ \vartheta_i = 0, &\quad \text{if } r_{R_i} < (1-\tau_i)r_{F_i} \end{aligned} \tag{19}$$

5.1.4 Agglomeration spillovers

Previous literature has found that agglomeration spillovers might shape the aggregate impact of new transit infrastructure (Tsivanidis, 2021). Hence, in this section I explain how I incorporate them into my quantitative framework. In particular, productivity per sector A_{js} can be decomposed into two components, one which is exogenous and captures fundamentals, and the other which is endogenous capturing externalities.

Productivity in each location depends on an exogenous component $\bar{A}_{js}$ that reflects the location's fundamentals (such as slope of the land, access to roads, etc.), and the endogenous employment in industry s in that location. I model endogenous productivity using constant scale elasticity, but allowing it to vary across sectors. In other words:

$$A_{js} = \bar{A}_{js}\left(\tilde{L}_{F_{js}}\right)^{\phi_{F_s}} \tag{20}$$

where ϕ_{F_s} measures the overall effect of agglomeration forces in sector s.[41]

5.1.5 Definition of Equilibrium

I assume that $A_{js} > 0$, $u_{ig} > 0$ and $\xi_{i\ell} > 0$ for all locations, sectors, and household groups. Corner solutions are taken care off in Appendix A2.

Definition 1. Given the model's parameters $\left\{\beta^k, \kappa^k, \theta^k, T^k, E^g, \eta^k, \alpha_{sg}, \sigma, \sigma_D, \phi_{Fs}, , \nu\right\}$, city population by household-group $\left\{N^k\right\}$, and exogenous location-specific characteristics $\left\{H_i, \xi_{i\ell}, \bar{u}_i^k, \bar{A}_{js}, t_{ij}, \tau_i\right\}$, the general equilibrium of the model is given by the vector $\left\{N_{Ri}^k, \mu_{i\ell}, L_{F_{jg}}, w_{jg}, r_{R_i}, r_{F_j}, \vartheta_i, \right\}$ such that labor markets clear, floorspace markets clear, and the population adds up to the city total.

The solution algorithm is presented in Appendix A2.

[41] To the extent that agglomeration economies are higher in male intensive industries, then agglomeration economies may limit the effect of improvements in the transit system on female earnings. Albeit this effect will ultimately depend on the city's geography and the kind of jobs women are gaining access to. The idea of sectoral agglomeration forces (also called external economies) is not new and recently papers have assessed their importance (e.g. Bartelme et al., 2019).

5.2 Calibration and Estimation From Aggregate Data

In the following table I specify which parameters I estimate from aggregate data and which ones I calibrate from the literature, and their respective values. In particular, I estimate from aggregate data the labor input share by industry-gender, and the household expenditures in housing (Panel A). Finally, I calibrate the elasticity of substitution between male and female workers in the production function, the elasticity of substitution of demand, and the labor input share by industry (Panel B).

[Table 4 here]

5.3 Estimation Outside the Model

5.3.1 Commuting elasticities

By taking logs to commuting probabilities in single and breadwinner households I can estimate the parameters of interest by PPML. From equations 9 and 13 I get the following :

$$\begin{aligned}
\log\left(\frac{L_{R^g_{ij}}}{N_{R^g_i}}\right) &= -\theta^g\kappa^g \cdot t_{ij} + \underbrace{\theta^g \cdot \log\left(w_{jg}\right)}_{\text{FE}_j} - \underbrace{\log \Psi_{R^g_i}}_{\text{FE}_i} \\
\log\left(\frac{L_{R^h_{ij,\ell=0}}}{N_{R^{mf}_{i,0}}}\right) &= -\theta^k\kappa^k \cdot t_{ij} + \underbrace{\theta^h \cdot \log\left(w_{jh}\right)}_{\text{FE}_j} - \underbrace{\log \Phi^h_{R_{i,\ell=0}}}_{\text{FE}_i}
\end{aligned}$$

which are exactly the same equations I estimate for Fact 1 reported in columns 1 to 4 of Table 1. I plug into the model the commuting elasticities I estimate for singles. For married households, I plug in the commuting elasticities of breadwinner households.[42] The underlying assumption is that among married of a given gender there is no heterogeneity in the shape parameter associated to their commuting preference shocks.[43] Finally, to get an estimate for the Frechet shape parameters in the commuting preferences, I follow the growing consensus in the literature (Ahlfeldt et al., 2015; Tsivanidis, 2021; Zarate, 2021) and set the iceberg disutility parameter κ^k to $\kappa^k = 0.01$, from which I get estimates of $\hat{\theta}^m = 5.14$, $\hat{\theta}^f = 6.17$, $\hat{\theta}^h = 4.66$, and $\hat{\theta}^w = 6.34$, as shown in Panel C of Table 4.

5.4 Estimation Within the Model

Model inversion. From observed data, I need to recover unobserved characteristics such as wages, productivities, amenities, household productivity, and land use wedges. We can recover values of location productivities, household productivities and amenities that rationalize the observed data as a model equilibrium. From the production-side we only need to observe workplace employment levels per industry rather than employment by gender. This is helpful since data on location-level employment

[42] While in the model I am not considering female breadwinner households as they make up less than 2% of households, I can still use data on them to estimate the commuting elasticity of married women.

[43] In part, I make this assumption because equations 12 and 18 state that conditioning on origin and destination fixed effects may not be enough to consistently estimate the commuting elasticities for dual-earner households. In any case, columns 5 to 6 show very similar elasticities as those shown in column 3 and 4.

by gender is unavailable in this setting. Intuitively, given male and female intensities per industry, a greater employment from a particular sector is informative of employment levels by gender. This is the same intuition as in Tsivanidis (2021). The following propositions formalize these ideas.

PROPOSITION 1: *Given data on residence by gender and household type, $N_{R^k_{i,\ell}}$, employment by gender $L_{F_{jg}}$ and commute costs d^k_{ij}, and parameters θ^k, there exists a vector of wages w_{jg} up to scale that rationalizes the observed data as a model equilibrium. Additionally, if one does not observe $L_{F_{jg}}$ but rather employment by industry $L_{F_{js}}$ and parameters α_{sg} and σ, there exists a vector of wages w_{jg} up to scale that rationalizes the observed data as a model equilibrium.*

Proof. See Appendix A3.1. □

Now we can turn to the problem of recovering the rest of unobservable variables.

PROPOSITION 2: *Given data on residence by gender and household type, $N_{R^k_{i,\ell}}$, employment by gender $L_{F_{jg}}$, rents r_{F_j},and commute costs d^k_{ij}, and parameters θ^k, α_s,σ_D there exists a vector of sales X_{js} and productivities A_{js} that rationalize the observed data as an equilibrium of the model. Additionally, we get the same result if instead of observing employment by gender we observe employment by industry $L_{F_{js}}$ and parameters σ, α_{sg}.*

Proof. See Appendix A3.2. □

PROPOSITION 3: *Given data on residence by gender and household type, $N_{R^k_{i,\ell}}$, employment by gender $L_{F_{jg}}$, available floorspace H_i, floorspace allocation ϑ_i, rents r_{R_i} and r_{F_j} and commute costs d^k_{ij}, and parameters θ^k, η^k, ν, σ^D, α_s, β^k, there exists a vector of amenities u^k_i, productivities A_{js}, household productivity $\xi_{i\ell=0}$, sales X_{js}, floorspace wedge τ_i, and total rents E that rationalize the observed data as an equilibrium of the model. Additionally, we get the same result if instead of observing employment by gender we observe employment by industry $L_{F_{js}}$ and parameters σ, α_{sg}.*

Proof. See Appendix A3.3. □

5.4.1 Agglomeration economies

By Proposition 2, I can recover productivity from $A_{js} = W^{\alpha_s}_{js} r^{1-\alpha_s}_{F_j} X^{1/(\sigma_D-1)}_{js} X$. Hence, productivity is the residual that ensures the definition for firm sales X_{js} to hold. In this way, we know that locations characterized by high employment levels relative to the price of commercial floorspace and wages are also locations with high productivity, at least according to the model. I use Proposition 2 to recover productivity in 2007, before transit improvements, and in 2017. Dividing by the geometric mean, taking logs, and then taking first differences, we can express our definition for agglomeration economies in equation 20 as:

$$\Delta \ln \frac{A_{js}}{\tilde{A}_s} = \phi_{F_s} \Delta \ln \frac{\tilde{L}_{F_{js}}}{\tilde{\tilde{L}}_F} + \Delta \ln \frac{\bar{A}_{js}}{\tilde{\bar{A}}_s}$$

Estimating the strength of agglomeration economies is a challenging endeavor as locations may be more productive because more people work there, or because locations with high productivity attract more workers. To identify this parameter, we need to use a shock to labor supply. Therefore, we can exploit the model and the introduction of new transit infrastructure to generate such an instrument. In particular, by taking a full differentiation to $\left\{\sum_k \sum_i \pi^k_{j|i,\ell} N_{R^k_i}\right\}$ with respect to commuting times from origin locations (while keeping market access terms constant) we get the following Bartik-style instruments:

$$Z^m_{Fj} = -\theta^m \kappa^m \sum_{i'} \frac{\pi^m_{j|i'} N_{R^m_{i'}}}{\sum_i \pi^m_{j|i} N_{R^m_i}} \cdot dt_{i'j} \qquad Z^f_{Fj} = -\theta^f \kappa^f \sum_{i'} \frac{\pi^f_{j|i'} N_{R^f_{i'}}}{\sum_i \pi^f_{j|i} N_{R^f_i}} \cdot dt_{i'j}$$

$$Z^w_{Fj} = -\theta^w \kappa^w \sum_{i'} \frac{\pi^w_{j|i',\ell=1} N_{R^{mf}_{i',\ell=1}}}{\sum_i \pi^w_{j|i,\ell=1} N_{R^{mf}_{i,\ell=1}}} \cdot dt_{i'j} \qquad Z^h_{Fj\ell=0} = -\theta^h \kappa^h \sum_{i'} \frac{\pi^h_{j|i',\ell=0} N_{R^h_{i',\ell=0}}}{\sum_i \pi^h_{j|i,\ell=0} N_{R^h_{i,\ell=0}}} \cdot dt_{i'j}$$

$$Z^h_{Fj\ell=1} = -\theta^h \kappa^h \sum_{i'} \frac{\pi^h_{j|i',\ell=1} N_{R^{mf}_{i',\ell=1}}}{\sum_i \pi^h_{j|i,\ell=1} N_{R^{mf}_{i,\ell=1}}} \cdot dt_{i'j} \quad - \quad \theta^w \kappa^w \sum_{i'} \frac{\pi^h_{j|i',\ell=1,j} \pi^w_{j|i',\ell=1} N_{R^{mf}_{i',\ell=1}}}{\sum_i \pi^h_{j|i,\ell=1} N_{R^{mf}_{i,\ell=1}}} \cdot dt_{i'j}$$

For each destination j, the Bartik-instruments weight the change in commuting times with the share of workers from each origin. I use inferred shares $\pi_{j|i}$ from 2007 data. Since I have many groups of households this calculation has to be made for each, providing us with five possible instruments to exploit. The identifying assumption is that the growth of exogenous productivities is uncorrelated with the Bartik instruments: $\mathcal{E}\left[\Delta \log \bar{A}_{js} \cdot Z^k_{Fj\ell}\right] = 0$.

To construct sector-specific instruments, I use the model to compute the share of female and male workers in each location-by-industry in 2007, which then I use to average the five instruments presented above in the following way:

$$Z_{Fjs} = \frac{L_{F_{jfs}}}{L_{Fjs}}(0.5 \cdot Z^f_{Fj} + 0.5 \cdot Z^w_{Fj}) + \frac{L_{F_{jms}}}{L_{Fjs}}(0.33 \cdot Z^m_{Fj} + 0.33 \cdot Z^h_{Fj\ell=1} + 0.33 \cdot Z^h_{Fj\ell=0})$$

So, in location-by-sectors that employ a greater share of females, my sector-specific instrument gives more weight to female Bartik shocks such as Z^f_{Fj} and Z^w_{Fj}, and vice-versa.

Results are shown in Panel A of Table 5.[44] Two observations are in order. First there is some heterogeneity in the strength of agglomeration economies across industries (see column 2 of Panel A). We have sectors such as Services with a scale elasticity statistically indistinguishable from zero, and Retail Trade where the scale elasticity is 0.24. Second, this heterogeneity might affect the extent to which the new infrastructure affects the gender gap since industries are located in different places of

[44] I control for a set of block-level covariates to account for differential growth in locations of such characteristics. In particular, I include the euclidean distance to the closest BRT station, the euclidean distance to the closest Metro station, and the euclidean distance to the Central Business District and its square. This allows me to control for anything that might have changed in places near stations but that are not related to improvement in commuting times. I also control for the physical size of the block and its square, its size in terms of the number of households in 2007, its slope, its elevation, and the number of dual-earner households as a share of married couples in 2007. Moreover, I include the average market access across household-types as measured in 2007. Finally, I control for the dependent variable at its value in 2007.

the city and employ different shares of female and male workers. If the new transit infrastructure provides women with access to female-intensive industries that have high agglomeration economies (such as textiles), then the gender gap might close compared to a counterfactual city where the new transit infrastructure is placed somewhere else.

[Table 5 here]

5.4.2 Labor supply elasticity

By Proposition 1, I can recover wages and compute market access measures $\Phi^{w}_{R_{i\ell=1}}$ and $\Phi^{h}_{R_{i\ell=0}}$. Then, by taking logs to the odds ratio of labor force participation in equation 14 and taking first differences I get:

$$\Delta \log \frac{\mu_i^{mf}}{1-\mu_i^{mf}} = \nu \left[\frac{\Delta \log \Phi_{R^w_{i,\ell=1}}}{\theta^w} - \frac{\Delta \log \Phi_{R^h_{i,\ell=0}}}{\theta^h} \right] + \nu \left(\Delta \log \xi_{i\ell=1} - \Delta \log \xi_{i\ell=0} \right)$$

where the term in bracket measures how the accessibility to jobs for dual-earner households improved relative to job accessibility for male breadwinner households. If changes in job access are correlated with changes in household productivity, then estimating this equation by OLS will lead to biased estimates of ν. To bypass this endogeneity concern I use the model to produce instruments. Since this is a supply-side parameter, we require variation from the demand. In particular, I use the TFP terms I recover from Proposition 2 to define the following instruments exploiting destination-level changes in TFP:

$$Z_{R^{mf}_{is}} = \sum_j \left(d^w_{ij} \right)^{-\theta^w} d \ln A_{js}$$

The identifying assumption is that $\mathbb{E}_i \left[\left(\Delta \log \xi_{i\ell=1} - \Delta \log \xi_{i\ell=0} \right) \cdot Z_{R^{mf}_{is}} \right] = 0$. Results are show in Panel B of Table 5.[45] I estimate a labor supply elasticity ν of 1.28.

5.4.3 Location elasticity

By Proposition 1 and 3, I can recover wages and household productivity and compute market access measures $\Psi_{R^g_i}$ and $\Psi_{R^{mf}_i}$. Manipulating the equations for location choices in 10 and 15, I get:[46]

[45]Once again I control for the complete set of euclidean distances (to the BRT, to the Metro, to the CBD). Moreover, I control for the physical size of the block, for its size in terms of the number of households in 2007, total employment in 2007, the number of dual-earner households as a share of married couples in 2007, and the block's slope and its area.

[46]Constants are defined as $\chi_g = \log \Delta \sum_l \left(\left(u^g_l \Psi^{1/\theta^{g,g}}_{R^{g,g}_l} \right)^{-1} r^{1-\beta^g}_{R_l} \right)^{-\eta^{g,g}}$ and $\chi_{mf} = \log \Delta \sum_l \left(\left(u^{mf}_l \Psi^{1/\nu}_{R^{mf}_l} \right)^{-1} r^{1-\beta^{mf}}_{R_l} \right)^{-\eta^{mf}}$.

$$
\begin{aligned}
\Delta \log \frac{N_{R_i^g}}{N^g} &= \chi_g + \eta^g \left[(\beta^g - 1)\, \Delta \log r_{R_i} + \frac{1}{\theta^g} \Delta \log \Psi_{R_i^g} \right] + \eta^g \Delta \log u_i^g \\
\Delta \log \frac{N_{R_i^{mf}}}{N^{mf}} &= \chi_{mf} + \eta^{mf} \left[\left(\beta^{mf} - 1 \right) \Delta \log r_{R_i} + \frac{1}{\nu} \Delta \log \Psi_{R_i^{mf}} \right] + \eta^{mf} \Delta \log u_i^{mf}
\end{aligned}
$$

If growth in exogenous amenities is correlated with the growth in market access measures $\Psi_{R_i^g}$ and $\Psi_{R_i^{mf}}$ then the estimation of the location elasticities η^g and η^{mf} would be biased. Once again, I exploit changes in TFP to construct instruments for the growth in market access measures. Results are shown in Panel B of Table 5.[47] My results indicate that singles ($\hat{\eta}^f \approx \hat{\eta}^m \approx 2$) are much more mobile than married couples ($\hat{\eta}^{mf} \approx 1$).

5.5 Over-identification and Commuting Probabilities

In this subsection, I examine the extent to which the model can account quantitatively for the observed variation in commuting probabilities. I use the Census data to compute the probability that a worker of a particular household type and gender commutes between any of the 49 districts of Lima in 2017, yielding 2401 pairs of bilateral commuting probabilities. I solve the model using 2017 exogenous characteristics and commuting times under two models, one in which commuting is independent and one where it depends on the other spouse's behavior.[48] Then, I compare the predicted commuting probabilities with those observed in the 2017 Census.

Figure B.6 plots a bin scatter of commuting probabilities in the model and in the data. Panel A focuses on females in dual-earner households. The left hand side of Panel shows that the model with independent commuting probabilities underestimate commuting flows. For every percentage point that the model predicts, the data display probabilities 13.45% lower since the slope is 0.8655. Moreover, the model is able to predict around 36% of the variation in the data. On the right hand side, Panel A of Figure B.6 reveals that on average the model with interdependent commuting probabilities correctly predicts commuting probabilities. In fact, the slope between predicted and observed commuting probabilities is 0.99. Moreover, the model is able to explain 52% of the variation in the data. Panel B exhibits the results for males in dual-earner households. The model with independence performs well since the slope between the commuting probabilities in the data and in the model is of 0.9779. However, the model with interdependent commuting probabilities is capable to improve performance, yielding a slope between probabilities in the data and in the model of 1.0142.

Therefore, despite the model being an abstraction, it is able to capture key features of couples' commuting patterns that the model with independent choices is unable to capture, especially in the case of females in dual-earner households.

[47] I include controls such as the euclidean distance to the Metro, to the BRT and to the CBD. I also include the size of the block, its slope and its elevation, the 2007 rents, and dependent variable at its value in 2007.

[48] Maximizing utility function: $U_{i\ell jj'}^{mf}(\omega) = u_i^{mf} \nu_i^{mf}(\omega) \cdot \xi_{i\ell}^{mf} \alpha_\ell^{mf}(\omega) \cdot \frac{\epsilon_j^h(\omega)}{d_{ij}^h} C_{i\ell j}^h(\omega)^{\beta^{mf}} H_{R_{i\ell j}^h}(\omega)^{1-\beta^{mf}} \cdot \left[\frac{\epsilon_{j'}^w(\omega)}{d_{ij'}^w} C_{i\ell j'}^w(\omega)^{\beta^{mf}} H_{R_{i\ell j'}^w}(\omega)^{1-\beta^{mf}} \right]^\ell$ subject to budget constraints $PC_{i\ell j}^h(\omega) + r_{R_i} H_{R_{i\ell j}^h}(\omega) = w_{jm}$ and $PC_{i\ell j'}^w(\omega) + r_{R_i} H_{R_{i\ell j'}^w}(\omega) = w_{j'f} \ell_i$ delivers a model with independent commuting probabilities.

6 The Impact of the Line 1 of the Metro and BRT

This section quantifies how the new transit infrastructure affected the gender earnings gap and aggregate real earnings through improved commute times. Moreover, I use the model to rationalize the reduced-form effects. It also provides evidence of the mechanisms in the paper, and how my quantitative results would change if I use a model in which choices are independent rather than interdependent.

I conduct a counterfactual exercise where I remove the Metro and the BRT from the city, while keeping exogenous characteristics fixed at their 2017 values. This informs about the impact of changing commuting times. I perform this counterfactual exercise for two cases, one in which spouses perform their commuting choices independently from each other, and one in which one spouse's choice is dependent on the other spouse's choice. To perform these counterfactual exercises, I use observed data to recover unobservables under the two models. Then, I solve each model conditionally on the corresponding unobservables.

For each gender in each type of household I construct a measure of real income across locations $RY_i^k = \frac{y_i^k}{P(\beta^k) r_R^{1-\beta^k}}$ where k indexes a gender-by-household group. Then, I compute aggregate real income RY^k simply by summing up RY^k over all locations. Since I am interested not only on aggregate income but on average earnings, I construct a measure of real income per worker $\bar{RY}^k = RY^k / N_{R^k}$, where N_{R^k} is the number of workers that belong to gender-household type k. Using these measures of real income I compute the gender earnings gap by dividing male real income by female real income and subtracting one: $\text{Gap} = RY^{k,m}/RY^{k,f} - 1$ or $\bar{\text{Gap}} = \bar{RY}^{k,m}/\bar{RY}^{k,f} - 1$, where now k indexes single and married households. Finally, I also compute the number of dual-earner households as a share of married households.

6.1 Impact on the gender earnings gap across locations

Figure 8 maps the evolution of real income RY_i and gender gap Gap_i^k across origins i caused by the improved times. Each color accounts for 20% of observations. First, the map in the left hand side of Panel A shows that real income increased mostly in the northern and southern fringe of the city, while it shrunk in central areas of the city and on the eastern fringe. This is the consequence of two features. (i) Households in the northern and southern fringe of the city were the main beneficiaries of the infrastructure, as it allowed them to commute faster to central areas. (ii) Increased labor supply coming from the northern and southern fringe pushed down wages, especially in destinations in central areas. In the right hand side of Panel A I exhibit the impact on the gender gap in aggregate real income. The most salient feature of this graph is that the gap decreased the least in central areas, and the most in the fringe. However, impacts vary by household-type and location. Panel B maps the impact on the gender gap in single households and in married households. Two differences between singles and married arise. First, gap reductions tend to be larger for married. Second, the decrease in the aggregate gender gap among married tend to be concentrated in the northern and southern fringe. This is also true for singles but to a lesser extent.

[Figure 8 here]

In Figure 9, I re-estimate my reduced-form specification comparing locations closer to stations against locations closer to yet-to-be-constructed locations to confirm the observations above. In Panel A, I show the impact on the aggregate real earnings gap per location, whereas in Panel B, I show the impact on the average real earnings gap. I show that both gaps decreased mainly in the fringe. This effect is almost entirely driven by married households for which aggregate gaps decreased by more than 13% and per worker gaps by 8.8%. Furthermore, the aggregate earnings gap increased for singles in some locations, but the average gap remained the same.

[Figure 9 here]

How different would the results be if, instead of implementing this paper's model, I use a model in which commute choices are independent within the couple? This is shown in Figure 10. Panel A reveals that married households experienced a decrease in the gap in aggregate earnings. However, Panel B shows that this is not due to a decrease in the gap in earnings per worker. Rather, the decrease shown in Panel A is driven by the fact that the improvement in commute times induced male breadwinner households to become dual-earner households, thereby increasing the mass of income accumulated by married women relative to married men.

[Figure 10 here]

In sum, the improvement in commute times reduced the gender gap among married households according to the model with interdependent choices. This effect was greater in the fringe of the city. Tacking this paper's question with a model with independent choices would have provided a very different result. Next, I summarize aggregate impacts.

6.2 Aggregate impact and decomposition

Table 6 reports my aggregate findings. It reports the percentage growth in the outcomes described above. Panel A shows the results of using the full model. First, both single males and single females see their income grow modestly by 0.60% and 0.39%. Since males aggregate income grew faster than females', the gender gap in aggregate earnings increased by 0.48%. When looking at average earnings, the gender gap increased by 0.56%. Second, turning to married households, aggregate income from male breadwinner households decreased by 1.61%, but average income increased by 0.66%. This is because many breadwinner households endogenously became dual-earner households. In fact, the new transit infrastructure led to an increase in female labor force participation of 1.74%, representing 9.48% of the total growth in labor force participation in the data. Since the number of dual-earner households is increasing endogenously, it is helpful to look at the increase on average earnings. Males' average earnings in dual-earner households grew by 1.17%, whereas females' average earnings rose by 2.19%. These impacts led to a reduction of the gender gap in aggregate and average earnings, which decreased by 4.13% and 2.54%, respectively. Finally, regarding the overall impact of the new transit infrastructure, aggregate and real income per worker grew by 1.07% and 0.70%.

For a set of commute times changes $dt_{ij'}, \forall i, j'$, I can use equation 7 to decompose the aggregate impact into three pieces: the direct use channel, the indirect use channel, and the general equilibrium impact (which is the difference between the total impact and the sum of the direct and the indirect use channels). Figure 11 depicts how the impact on aggregate income can be decomposed into these three pieces for singles, dual-earner households, and women in dual-earner households against men in male-breadwinner households. Panel A does so when using the model with interdependent choices. It reveals that from a direct use perspective, the Metro and the BRT marginally reduced the gap among singles. In other words, single women tended to use the routes that were improved marginally more than single men. However, general equilibrium effects favored single men, increasing marginally the gap among singles.[49] Why general equilibrium effects favored men relative to women? In part this is because the increased labor supply of women led to a decrease in average wages at workplaces. On average, wages decreased by 1.6% for males, they decreased by 2.3% for females.

Panel A also shows the impacts among spouses in dual-earner households. It reveals that the direct use channel contributed the most in reducing gaps. This means that married women were initially more exposed to the routes that were improved. However, the indirect use channel and the general equilibrium effects tended to widen the gaps. Recall that according to Equation 7, the probability that wives will commute to low-wage locations increases through the indirect use channel when the initial conditional commuting probabilities to these locations are high enough. Since women tended to commute to low-wage locations, the indirect channel led to a widening of the gap.

Finally, this figure also shows the comparison of women in dual-earner households against men in male breadwinner households. Here, the general equilibrium channel tended to decrease the gap in aggregate earnings. This is because the share of male breadwinner households decreased, and so the mass of income they generated also decreased relative to dual-earner households.

Panel B of Table 6 shows the results when using a model with independent commute choices. Three observations are in order. First, this model overestimates the impact of the new transit infrastructure on aggregate real income by $1 - 0.0107/0.0137 = 22\%$ and on real income per worker by $1 - 0.0070/0.0105 = 33\%$. Second, while the gender gap in aggregate earnings in married households closes by 2.01%, this is mainly due to the labor force participation margin leading to an overall increase in the mass of income accumulated by females. When looking at earnings per worker, this model suggests that the gap in married households increased by 0.23% instead of shrinking by 2.54%. Third, this model underestimates the impact on the gender gap in aggregate earnings by $1 - 0.0128/0.0215 = 40\%$, and the gender gap in earnings per worker by $1 - (0.0008/0.0105) = 92\%$, leading to a switch in the sign.

Panel B of Figure 11 shows the decomposition of the effects into two channels: the direct use channel and the general equilibrium channel. First, the direct use channel, as inferred by the two models, should not be expected to be the same because it depends on commute probabilities, changes in wages and geography, and changes in commute time. While commute time changes are the same in both models as they are data, wages and probabilities are estimated based on the structure of the models and the observed data. Second, results for singles in both models are qualitatively similar, with the general equilibrium channel tending to increase the gap and the direct use channel tending to decrease it. Third, among dual-earner households, the results are also quantitatively similar to those

[49] The indirect use channel is absent for singles.

for singles. However, in the model with interdependent choices, the results for dual-earner households are stronger than for singles. Finally, when comparing women in dual-earner households to men in male breadwinner households, the gap decreased mainly due to general equilibrium effects, which are entirely due to the labor force participation choice. The direct use channel plays no role in this comparison.

All in all, allowing for interdependent commute choices is quantitatively important when evaluating not only the impact of the new transit infrastructure on the gender gap and overall efficiency, but also when understanding the margins through which infrastructure affects households. In fact, this exercise confirms that the gender gap decreased for married households because the new transit infrastructure improved the routes that married women were using the most. Without accurate data on commute probabilities, the model with independent choices would not have been able to accurately measure married women's exposure to the new infrastructure.

[Table 6 here]

[Figure 11 here]

6.3 Mechanisms

In this section I turn off mechanisms from the full model in order to assess their quantitative relevance. Results are shown in Table B.3. For comparison, the first column shows the aggregate impact on real income and the gender gap using the complete model.

The second column removes externalities, and assume that productivity and amenity terms observed in 2017 remain fixed. This column reveals that the impact on real income and real income per worker rises to 1.39% and 1.02%, respectively. This negative effect coming from externalities arises despite that productivity increased in every sector as a result of the new transit infrastructure. According to the full model, average productivity across the seven sectors grew by 0.53%. However, the price index markedly decreased in the counterfactual city with no externalities. In the full model, the price index decreased by 1.78%, whereas with no externalities it decreased by 3.02%. The explanation is that agglomeration externalities make the labor demand more elastic. Then, shifts in the labor supply have milder effects on wages, which prevents the price index to fall. In fact, males and female wages at destinations decreased by 1.61% and 2.28% in the full model, whereas they decreased by 2.52% and 3.17% in the model without externalities. Finally, the impact on the aggregate gender gap remains around of -2.00%, although the impact on the gender gap in earnings per worker becomes -0.84% rather than -1.05% as it was in the full model. This is driven by single workers, for which the impact on the gender gap on average earnings almost doubles from 0.48% to 0.89%.

The third column keeps the distribution of households across locations fixed at its 2017 value. In other words, it assumes that households cannot reallocate in response to the removal of the transit infrastructure. Compared to column (1), the impacts on real income and on real income per capita are 42% and 79% lower, which suggest that mobility of households is an important mechanism. Furthermore, while the impact on the aggregate gender gap remains quite constant, the impact on the gender gap in terms of average earnings becomes -0.78% instead of -1.05%. This is because, by keeping single households at their initial location, in this counterfactual exercise, it turns out single males are the main beneficiaries of the new transit infrastructure. This is reflected on the impact on the average gender gap, which becomes 0.98% instead of 0.56%.

Column (4) assumes that married couples cannot endogenously change their labor force participation. This reduces the impact on aggregate real income by $1 - 0.0096/0.0107 = 10\%$ which speaks to the moderate importance of this margin in the aggregate. Moreover, preventing households to change their labor participation choice has some distributional consequences through the general equilibrium. In particular, when this margin is activated, increased supply in female labor tends to decrease females wages relative to male wages. This is reflected on the gender gap in earnings per worker. When married couples have the possibility of changing their labor supply choice, the overall gender gap decreases by 1.05% as a consequence of the new transit infrastructure. Preventing this choice also prevents female wages to decrease even further, leading to a steeper shrinking of the gender gap in earnings per worker, as it decreases by 1.59%.

Finally, Column (5) turns off the three mechanisms above. Doing so decreases the growth in aggregate real income from 1.07% to 0.75%. So, between the three they explain 30% of the growth. However, earnings per worker remains almost unchanged. While the mobility margin is an important mechanism through which improved commute times increase real income per worker, the LFP and externalities tended to have the opposite effect. Turning to the gender gap, without the three mechanisms, the reduction in the aggregate earnings gap becomes 1.14% instead of 2.15%. However, the reduction in gender gap in earnings per worker increases from 1.05% to 1.49%. This is mainly driven by the LFP margin. Increased female supply decreases average women's average earning since they labor demand is downward sloping.

In sum, this analysis suggests that externalities, household mobility and the labor force participation choice shape the aggregate impact in important ways.

6.4 Understanding the reduced-form results

In the reduced-form section of the paper, I explained that, while location choices might be indicative of overall welfare, they conflate changes in market access, changes in rents, and changes in amenities. In this section, I open up this box and re-estimate equation 3 in a differences-in-differences framework using location outcomes from the counterfactual exercise. These capture the effect of the new transit infrastructure that goes through commuting times (including rent increases generated due to increased commuting times), but not the impact that goes through exogenous characteristics changing. Thus, I also estimate equation 3 using backed-out exogenous amenities from 2007 and 2017 as dependent variable.

Results are shown in Figure B.7. These results show that amenities worsened in locations near to station. This explains apparent divergence between the counterfactual exercise and the data in terms of total population. While better commuting times fostered households to move closer to stations, amenities pushed households out. For completeness, I also show this exercise for other outcomes and a similar pattern emerges: exogenous amenities for singles and married changed in non-negligible ways.[50]

In Figure B.8, I estimate the heterogeneous impacts by distance to the CBD. These results reveal that amenities worsened especially in areas farther away from the CBD. Furthermore, relative to married households, amenities for singles improved in the fringe, although data and counterfactual move similarly. When comparing male amenities to female amenities, on average they were similarly affected across locations, explaining why the counterfactual impacts follow a similar pattern as the impacts on the data. A similar result emerges for the odds ratio in wives' labor force participation.

7 Policy Counterfactual: Completing the Metro's Network

In this section, I use the model to evaluate what would happen if the Lines 2 and 4 of the Metro are completed. To do so, I assume that speeds in these lines are similar to those of Line 1. I then keep 2017 exogenous characteristics constant and incorporate Lines 2 and 4 to the city. I compare this counterfactual scenario against (a) a scenario where the Line 1 of the Metro and the BRT are active, and (b) a scenario where neither Metro Lines or the BRT are active. The first comparison informs us about the incremental impact of completing the network, whereas the second comparison informs us about the overall impact of the Metro and the BRT.

Results are shown in Table 7. Panel A reveals that on average real income would grow an additional 0.95% compared to the situation where only the Line 1 and the BRT are active. Also, income per worker would increase by 0.56%. Moreover, the gap would decrease by 1.87%. The gap in average earnings would shrink by 0.56%. The effect on the gap is driven by married couples. For these households, the gender earnings gap would decrease by 4.05%. In terms of income per worker, the gap would decrease by 2.26%. Overall, impacts are comparable in magnitude and sign to those of the Line 1 of the Metro and the BRT. The main difference is that in this case, the gap among singles increases in a non-negligible way, i.e. by 1.12%, or 1.31% if looking at the gap per worker.

Panel B exhibits the overall impact of the Metro and the BRT. The complete Metro network plus the BRT increases real income by 2.03%, and average earnings per worker by 1.26%. Also, the gap decreases by 3.98% on aggregate. The gap in income per worker decreases by 1.60%. Unsurprisingly, this is due to a decrease in the gap among married. In these households, the gap in aggregate earnings decreases by 8.01%, whereas the gap in earnings per worker decreases by 4.75%. However, among singles the gap increases by 1.61%. The gap in earnings per worker increases by 1.88%. Appendix B.9 shows the decomposition of these aggregate impacts. Compared to the analysis of the impact of the Line 1 and the BRT, I observe similar patterns. Finally, in an unreported analysis I use the simulated data to study the reduced-form impact across locations. Once again, I find that impacts on the gap are concentrated on areas closer to remote stations.

[Table 7 here]

[50] This is not accounted for in the counterfactual exercise given above. The reason is that I focus on how changes in commute times alone can affect married households. However, it is true that new transit infrastructure can also affect local characteristics as in Brinkman and Lin (2022).

8 Conclusion

I have uncovered two new insights about married households and their commuting choices. The first insight is that commuting choices are interdependent across spouses. In particular, as the wage of one of the spouses increases and secures a higher standard of living, households are willing to give up more of the other spouses' income to decrease commute costs. Thus, improving commute times can affect households through two channels: the direct use channel and the indirect use channel. The second insight is that commuting flows follow a gravity structure, making the model tractable and easier to estimate.

My main results show that the model with interdependent commuting choices fits better the data on bilateral commuting flows of dual-earner households than the model without interdependence. They also show that reduced commute times can reduce the gender gap, especially in places were access increased the most.

Policymakers should take into account the interdependence of commuting choices within households when designing and implementing transit infrastructure projects in order to maximize the positive impact on reducing the gender earnings gap and improving overall economic outcomes. For example, governments could prioritize improving infrastructure and services in less developed areas of cities, which tend to be farther from job opportunities. However, it's important to note that through the indirect use channel in commuting choices, improving one spouse's prospects could negatively affect the other spouse's earnings.

References

[1] Acemoglu, D., Autor, D. H., and Lyle, D. (2004). "Women, war, and wages: The effect of female labor supply on the wage structure at midcentury". In: *Journal of political Economy* 112.3, pp. 497–551.

[2] Adao, R., Arkolakis, C., Esposito, F., et al. (2020). *General equilibrium effects in space: Theory and measurement*. Tech. rep.

[3] Adao, R., Costinot, A., and Donaldson, D. (2017). "Nonparametric counterfactual predictions in neoclassical models of international trade". In: *American Economic Review* 107.3, pp. 633–89.

[4] Adda, J., Dustmann, C., and Stevens, K. (2017). "The career costs of children". In: *Journal of Political Economy* 125.2, pp. 293–337.

[5] Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). "The economics of density: Evidence from the Berlin Wall". In: *Econometrica* 83.6, pp. 2127–2189.

[6] Alam, M. M., Cropper, M., Herrera Dappe, M., and Suri, P. (2021). "Closing the Gap". In:

[7] Allen, T. and Arkolakis, C. (2014). "Trade and the Topography of the Spatial Economy". In: *The Quarterly Journal of Economics* 129.3, pp. 1085–1140.

[8] — (2020). *The welfare effects of transportation infrastructure improvements*. Tech. rep.

[9] Allen, T., Arkolakis, C., and Li, X. (2015). "Optimal city structure". In: *Yale University, mimeograph*.

[10] Alon, T., Coskun, S., Doepke, M., Koll, D., and Tertilt, M. (2022). "From mancession to shecession: Women's employment in regular and pandemic recessions". In: *NBER Macroeconomics Annual* 36.1, pp. 83–151.

[11] Angelov, N., Johansson, P., and Lindahl, E. (2016). "Parenthood and the gender gap in pay". In: *Journal of Labor Economics* 34.3, pp. 545–579.

[12] Bardóczy, B. (2022). "Spousal insurance and the amplification of business cycles". In: *Unpublished Manuscript, Northwestern University*.

[13] Bartelme, D. G., Costinot, A., Donaldson, D., and Rodriguez-Clare, A. (2019). *The textbook case for industrial policy: Theory meets data*. Tech. rep. National Bureau of Economic Research.

[14] Bartelme, D. (2018). "Trade Costs and Economic Geography: Evidence from the U.S." In:

[15] Becker, G. S. (1985). "Human capital, effort, and the sexual division of labor". In: *Journal of Labor Economics* 3.1, Part 2, S33–S58.

[16] Behrens, K., Mion, G., Murata, Y., and Suedekum, J. (2017). "Spatial frictions". In: *Journal of Urban Economics* 97, pp. 40–70.

[17] Black, D. A., Kolesnikova, N., and Taylor, L. J. (2014). "Why do so few women work in New York (and so many in Minneapolis)? Labor supply of married women across US cities". In: *Journal of Urban Economics* 79, pp. 59–71.

[18] Bryan, G. and Morten, M. (2019). "The aggregate productivity effects of internal migration: Evidence from Indonesia". In: *Journal of Political Economy* 127.5, pp. 2229–2268.

[19] Carta, F. and De Philippis, M. (2018). "You've come a long way, baby. Husbands' commuting time and family labour supply". In: *Regional Science and Urban Economics* 69, pp. 25–37.

[20] Cha, Y. and Weeden, K. A. (2014). "Overwork and the slow convergence in the gender gap in wages". In: *American Sociological Review* 79.3, pp. 457–484.

[21] Chiappori, P.-A., Giménez-Nadal, J. I., Molina, J. A., and Velilla, J. (2022). "Household Labor Supply: Collective Evidence in Developed Countries". In: *Handbook of Labor, Human Resources and Population Economics*, pp. 1–28.

[22] Clark, W. A., Huang, Y., and Withers, S. (2003). "Does commuting distance matter?: Commuting tolerance and residential change". In: *Regional Science and Urban Economics* 33.2, pp. 199–221.

[23] Combes, P.-P., Duranton, G., and Gobillon, L. (2008). "Spatial wage disparities: Sorting matters!" In: *Journal of Urban Economics* 63.2, pp. 723–742.

[24] Cubas, G., Juhn, C., and Silos, P. (2019). *Coordinated work schedules and the gender wage gap.* Tech. rep. National Bureau of Economic Research.

[25] Doepke, M. and Tertilt, M. (2016). "Families in macroeconomics". In: *Handbook of macroeconomics*. Vol. 2. Elsevier, pp. 1789–1891.

[26] Fajgelbaum, P. D. and Schaal, E. (2020). "Optimal transport networks in spatial equilibrium". In: *Econometrica* 88.4, pp. 1411–1452.

[27] Farré, L., Jofre-Monseny, J., and Torrecillas, J. (2020). "Commuting time and the gender gap in labor market participation". In:

[28] Feenstra, R. C., Luck, P., Obstfeld, M., and Russ, K. N. (2018). "In search of the Armington elasticity". In: *Review of Economics and Statistics* 100.1, pp. 135–150.

[29] Field, E. and Vyborny, K. (2022). "Women's Mobility and Labor Supply: Experimental Evidence from Pakistan". In: *Asian Development Bank Economics Working Paper Series* 655.

[30] Goldin, C. (2014). "A grand gender convergence: Its last chapter". In: *American Economic Review* 104.4, pp. 1091–1119.

[31] Greenstone, M., Hornbeck, R., and Moretti, E. (2010). "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings". In: *Journal of Political Economy* 118.3, pp. 536–598.

[32] Greenwood, J., Guner, N., and Marto, R. (2021). *The Great Transition: Kuznets Facts for Family-Economists.* Tech. rep. National Bureau of Economic Research.

[33] Gu, Y., Guo, N., Wu, J., and Zou, B. (2021). "Home Location Choices and the Gender Commute Gap". In: *Journal of Human Resources*, 1020–11263R2.

[34] Gutierrez, F. (2018). "Commuting Patterns, the Spatial Distribution of Jobs and the Gender Pay Gap in the US". In: *Available at SSRN 3290650.*

[35] Heblich, S., Redding, S. J., and Sturm, D. M. (2020). "The making of the modern metropolis: evidence from London". In: *The Quarterly Journal of Economics* 135.4, pp. 2059–2133.

[36] ILO (2017). "World Employment and Social Outlook: Trends for women 2017". In: *International Labour Office – Geneva: ILO.*

[37] JICA (2013). *Data Collection Survey on Urban Transport for Lima and Callao Metropolitan Area, Final Report.* Tech. rep. https://openjicareport.jica.go.jp/pdf/12087516_01.pdf. Accessed January 2020.

[38] Johnson, M. and Keane, M. P. (2013). "A dynamic equilibrium model of the US wage structure, 1968–1996". In: *Journal of Labor Economics* 31.1, pp. 1–49.

[39] Kawabata, M. and Abe, Y. (2018). "Intra-metropolitan spatial patterns of female labor force participation and commute times in Tokyo". In: *Regional Science and Urban Economics* 68, pp. 291–303.

[40] Kleven, H., Landais, C., and Søgaard, J. E. (2019). "Children and gender inequality: Evidence from Denmark". In: *American Economic Journal: Applied Economics* 11.4, pp. 181–209.

[41] Kline, P. and Moretti, E. (2014). "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority". In: *The Quarterly Journal of Economics* 129.1, pp. 275–331.

[42] Kwon, E. (2020). *Why Do Improvements in Transportation Infrastructure Reduce the Gender Gap in South Korea?* Tech. rep.

[43] Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). "Gender differences in job search: Trading off commute against wage". In: *The Quarterly Journal of Economics* 136.1, pp. 381–426.

[44] Liu, S. and Su, Y. (2020). "The geography of jobs and the gender wage gap". In: *The Review of Economics and Statistics*, pp. 1–27.

[45] Lundborg, P., Plug, E., and Rasmussen, A. W. (2017). "Can women have children and a career? IV evidence from IVF treatments". In: *American Economic Review* 107.6, pp. 1611–37.

[46] Mankart, J. and Oikonomou, R. (2017). "Household search and the aggregate labour market". In: *The Review of Economic Studies* 84.4, pp. 1735–1788.

[47] Mas, A. and Pallais, A. (2020). "Alternative work arrangements". In: *Annual Review of Economics* 12, pp. 631–658.

[48] Medina, C., Nyshadham, A., Ramos, D., Tamayo, J., and Tiew, A. (2021). "Spatial Mobility, Economic Opportunity, and Crime". In:

[49] Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). "Commuting, migration, and local employment elasticities". In: *American Economic Review* 108.12, pp. 3855–90.

[50] 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". In: *Sustainability* 11.10, p. 2795.

[51] Owens III, R., Rossi-Hansberg, E., and Sarte, P.-D. (2020). "Rethinking Detroit". In: *American Economic Journal: Economic Policy* 12.2, pp. 258–305.

[52] Reuben, E., Wiswall, M., and Zafar, B. (2017). "Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender". In: *The Economic Journal* 127.604, pp. 2153–2186.

[53] Roca, J. D. L. and Puga, D. (2017). "Learning by working in big cities". In: *The Review of Economic Studies* 84.1, pp. 106–142.

[54] Rosenthal, S. S. and Strange, W. C. (2004). "Evidence on the nature and sources of agglomeration economies". In: *Handbook of regional and urban economics*. Vol. 4. Elsevier, pp. 2119–2171.

[55] — (2012). "Female entrepreneurship, agglomeration, and a new spatial mismatch". In: *Review of Economics and Statistics* 94.3, pp. 764–788.

[56] Severen, C. (2021). “Commuting, labor, and housing market effects of mass transportation: Welfare and identification”. In:

[57] Tsivanidis, N. (2021). “Evaluating the impact of urban transit infrastructure: Evidence from Bogotá’s TransMilenio”. In:

[58] Zárate, R. D. (2021). “Factor Allocation, Informality and Transit Improvements: Evidence from Mexico City.” In:

Table 1: Commuting Elasticity

	Single HH		Breadwinner HH		Dual-Earner HH	
	Males	Females	Males	Females	Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Time	-0.0514	-0.0617	-0.0466	-0.0634	-0.0480	-0.0673
	(0.0034)***	(0.0038)***	(0.0033)***	(0.0041)***	(0.0032)***	(0.0041)***
Origin FE	X	X	X	X	X	X
Destination FE	X	X	X	X	X	X
Gap (Female/Male-1)		20%		36%		40%
N	2500	2500	2500	2500	2500	2500

Table 2: Reduced Form Impact of New Transit Infrastructure on Rents and Income

	A+H Rents	HH Income	HH Exp.	Poverty
	(1)	(2)	(3)	(4)
$Treat_l^M \times Before2009_t$	-0.0010	0.0311	0.0687	-0.0152
	(0.0698)	(0.0781)	(0.0468)	(0.0325)
$Treat_l^M \times After2009_t$	-0.0216	0.0809	0.1059	-0.0496
	(0.0671)	(0.0651)	(0.0464)**	(0.0275)*
$Treat_l^{BRT} \times Before2009_t$	-0.0229	0.0376	-0.0651	0.0448
	(0.0830)	(0.0937)	(0.0516)	(0.0307)
$Treat_l^{BRT} \times After2009_t$	-0.0414	0.1052	-0.0183	-0.0380
	(0.0686)	(0.0900)	(0.0486)	(0.0249)
Block FE	X	X	X	X
Controls	X	X	X	X
N	24246	24246	24246	24246
Est. Method	PPML	PPML	PPML	OLS

Table 3: Time-Event Study of New Transit Infrastructure on Census Outcomes

Panel A:	Total Population	Singles to Married	Females to Males
	(1)	(2)	(3)
$Treat \times (Year = 1993)$	0.0938	0.0396	0.0234
	(0.1070)	(0.0444)	(0.0256)
$Treat \times (Year = 2017)$	-0.1242	0.0812	-0.0033
	(0.0354)***	(0.0276)***	(0.0063)
Panel B:	Single to Married (Males)	Single to Married (Females)	Dual-Earners to Male Breadwinner
	(1)	(2)	(3)
$Treat \times (Year = 1993)$	0.0287	0.0461	-0.0073
	(0.0502)	(0.0442)	(0.0491)
$Treat \times (Year = 2017)$	0.0894	0.0764	-0.0516
	(0.0298)***	(0.0288)***	(0.0315)
Block FE	X	X	X
Year FE	X	X	X
N	375	375	375

Table 4: Estimation from Aggregate Data, Calibration, and Gravity Equation

Parameter	Description	Value	Source
Panel A: Estimation from Aggregate Data			
$[\alpha_{f1}, \alpha_{f2}, \alpha_{f3}, \alpha_{f4}, \alpha_{f5}, \alpha_{f6}, \alpha_{f7}]$	Female input share by industry	$[0.21, 0.42, 0.41, 0.35, 0.29, 0.29, 0.20]$	ENAHO
$[1-\beta_m, 1-\beta_f, 1-\beta_{mf}]$	Household expenditure in housing	$[0.212, 0.208, 0.171]$	ENAHO
Panel B: Calibration			
Parameter	Description	Value	Source
σ	Male-female elasticity of substitution	2	Johnson and Keane (2013)
σ_D	Elasticity of substitution of demand	5	Freenstra et al. (2018)
α_s	Labor input share by industry	1-0.2 $\forall s$	Ahlfeldt et al. (2015)
Panel C: Commuting Elasticity			
Parameter	Description	Value	Source
$[\kappa^m, \kappa^f, \kappa^h, \kappa^w]$	Iceberg disutility	$[0.01, 0.01, 0.01, 0.01]$	Tsivanidis (2021)
$[\hat{\theta}^m, \hat{\theta}^f, \hat{\theta}^h, \hat{\theta}^w]$	Shape parameter in commuting preferences	$[5.14, 6.17, 4.66, 6.34]$	Gravity Equation. See Section 5.3.1.

Considered industries are (in order): Manufacture (w/o textiles), Textiles, Services, Business Services, Wholesale Trade, Retail Trade, Transportation. ENAHO stands for the 2007-2018 waves of the national household survey.

Table 5: Estimation of Remaining Parameters

Panel A: Agglomeration Externalities (ϕ_{F_s})

	OLS	2SLS	F - Weak Id
	(1)	(2)	(3)
Manufacture (w/o textiles)	0.6230	0.1057	43.54
	(0.0146)***	(0.0877)	
Manufacture (textiles)	0.4836	0.1668	19.02
	(0.0184)***	(0.0968)*	
Services	0.7255	0.0590	19.72
	(0.0126)***	(0.1561)	
Business Services	0.5991	0.1837	30.98
	(0.0200)***	(0.0888)**	
Wholesale Trade	0.6261	0.1591	39.43
	(0.0158)***	(0.0814)*	
Retail Trade	0.8204	0.2435	30.09
	(0.0126)***	(0.0924)***	
Transportation	0.6112	0.1573	42.37
	(0.0138)***	(0.0798)**	

Panel B: Labor Supply and Location Elasticities

	OLS	2SLS	F - Weak Id
	(1)	(2)	(3)
Labor Supply Elasticity (ν)	0.3644	1.2809	57.84
	(0.0776)***	(0.2418)***	
Location Single Females (η^f)	0.1801	2.0458	30.47
	(0.0707)***	(0.4363)***	
Location Single Males (η^m)	0.2046	1.8154	31.13
	(0.0733)**	(0.4121)***	
Location Married (η^{mf})	0.1106	1.0155	42.71
	(0.0519)***	(0.1853)***	

Table 6: The Aggregate Impact of the Line 1 of the Metro and the BRT

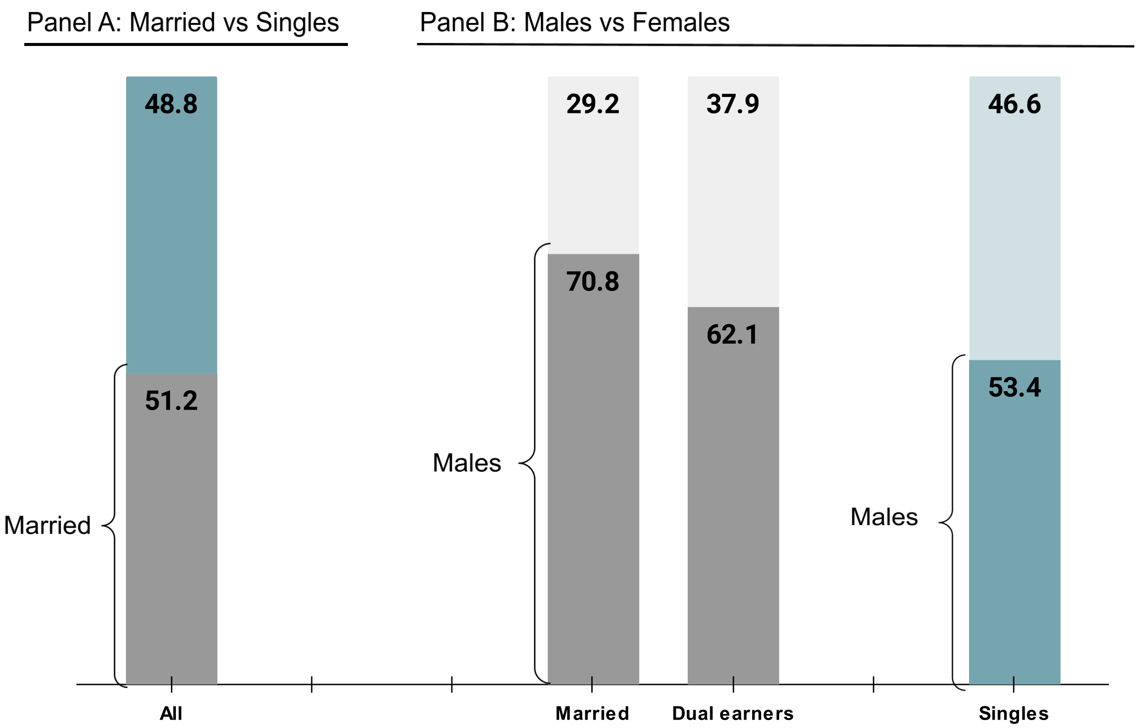
Panel A: Interdependent Commuting Choices								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0060	0.0039	-0.0161	0.0292	0.0396	0.0070	0.0213	0.0107
Real Income per Worker	0.0060	0.0039	0.0066	0.0117	0.0219	0.0070	0.0122	0.0070
Gender Gap	0.0048			-0.0413		-0.0215		
Gender Gap (per Worker)	0.0056			-0.0254		-0.0105		
# D.E. / # Married				0.0174				
Panel B: Independent Commuting Choices								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0097	0.0097	-0.0072	0.0273	0.0261	0.0114	0.0196	0.0137
Real Income per Worker	0.0097	0.0097	0.0123	0.0124	0.0112	0.0114	0.0118	0.0105
Gender Gap	0.0000			-0.0201		-0.0128		
Gender Gap (per Worker)	0.0000			0.0023		-0.0008		
# D.E. / # Married				0.0148				

Table 7: The Aggregate Impact of Completing the Metro Network

Panel A: Incremental Impact of Metro Lines 2 and 4								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0039	-0.0011	-0.0175	0.0295	0.0386	0.0063	0.0185	0.0095
Real Income per Worker	0.0039	-0.0011	0.0072	0.0112	0.0201	0.0063	0.0090	0.0056
Gender Gap	0.0112			-0.0405		-0.0187		
Gender Gap (per Worker)	0.0131			-0.0226		-0.0056		
# D.E. / # Married				0.0181				

Panel B: Overall Impact of Metro Lines (1,2, 4) and the BRT								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0099	0.0028	-0.0333	0.0596	0.0797	0.0133	0.0402	0.0203
Real Income per Worker	0.0099	0.0028	0.0139	0.0230	0.0424	0.0133	0.0213	0.0126
Gender Gap	0.0161			-0.0801		-0.0398		
Gender Gap (per Worker)	0.0188			-0.0475		-0.0160		
# D.E. / # Married				0.0358				

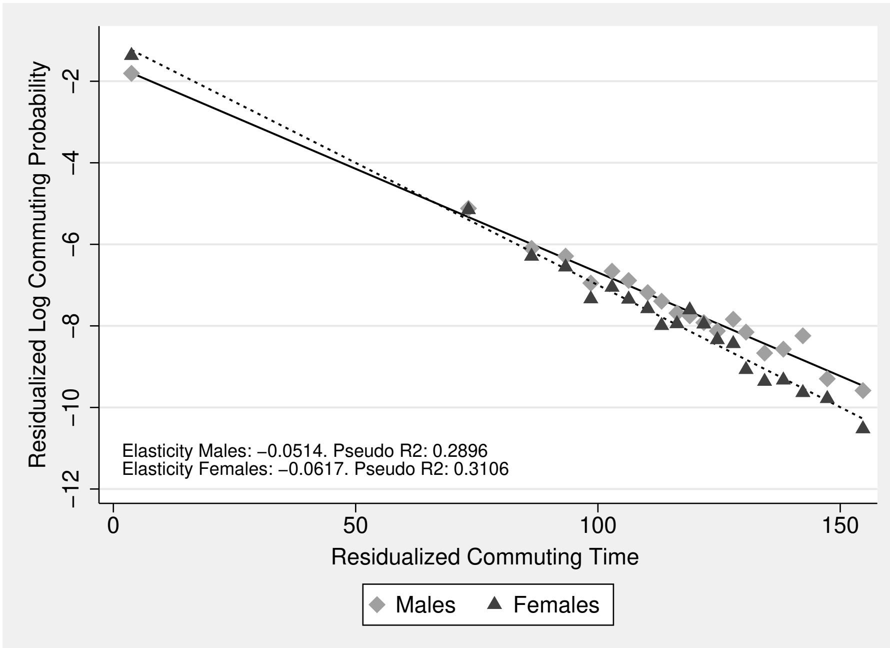
Figure 1: Distribution of Time Spent in Commuting Across Genders and Civil Status (%)



Notes: Households were classified according to definitions in Section A1.

Figure 2: Gravity in Commuting Flows among Singles and Married

Panel A: Singles



Panel B: Married

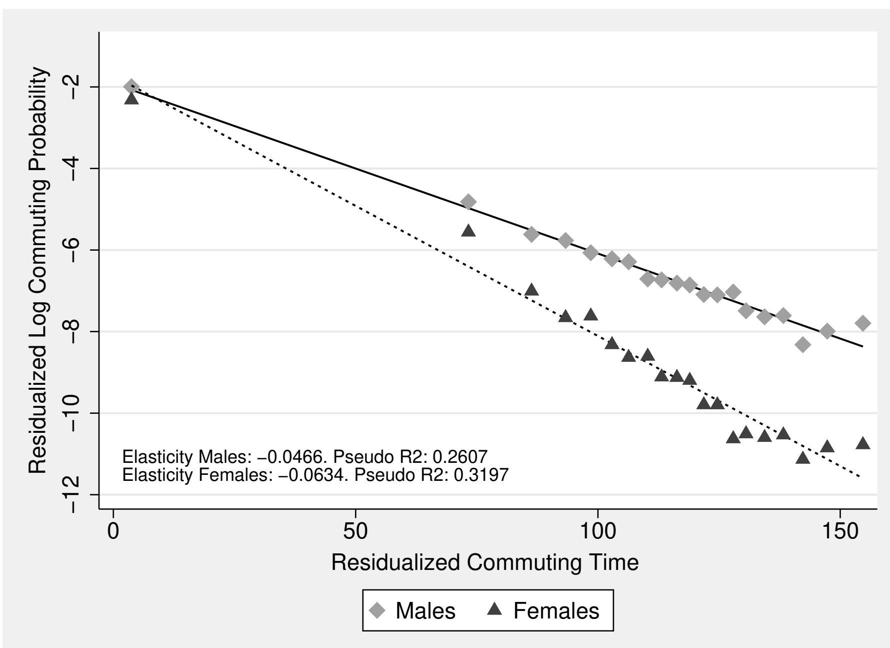
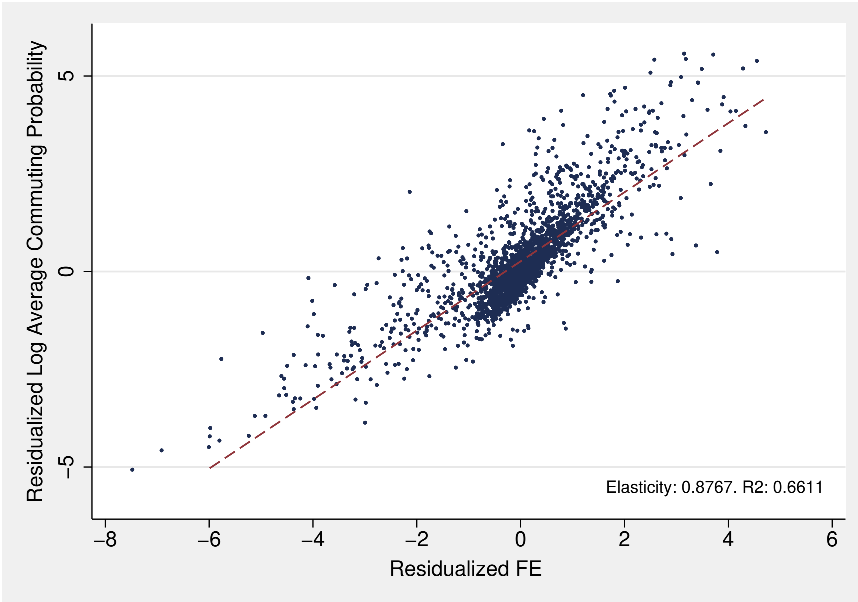


Figure 3: Interdependence in Commuting Choices within Dual-earner Households

Panel A: Husband



Panel B: Wife

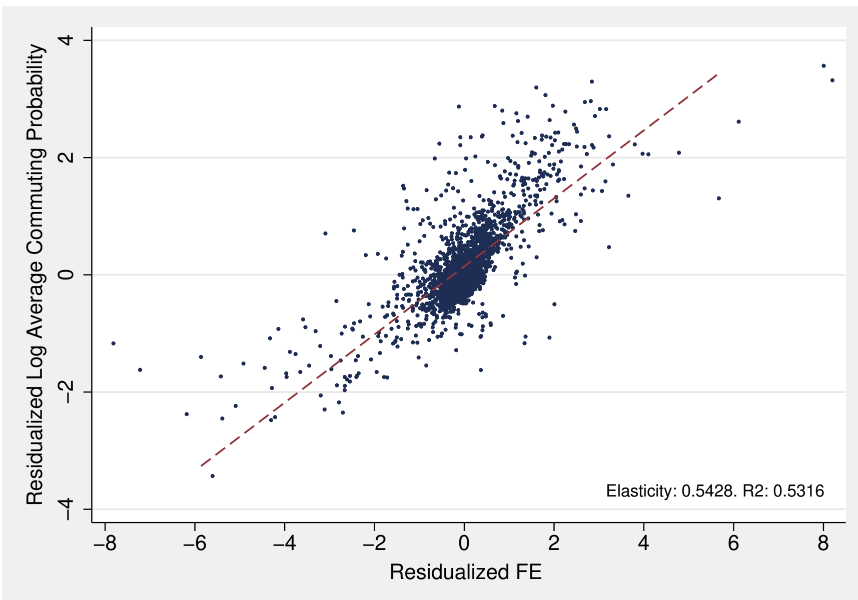


Figure 4: Constructed and Planned Lines

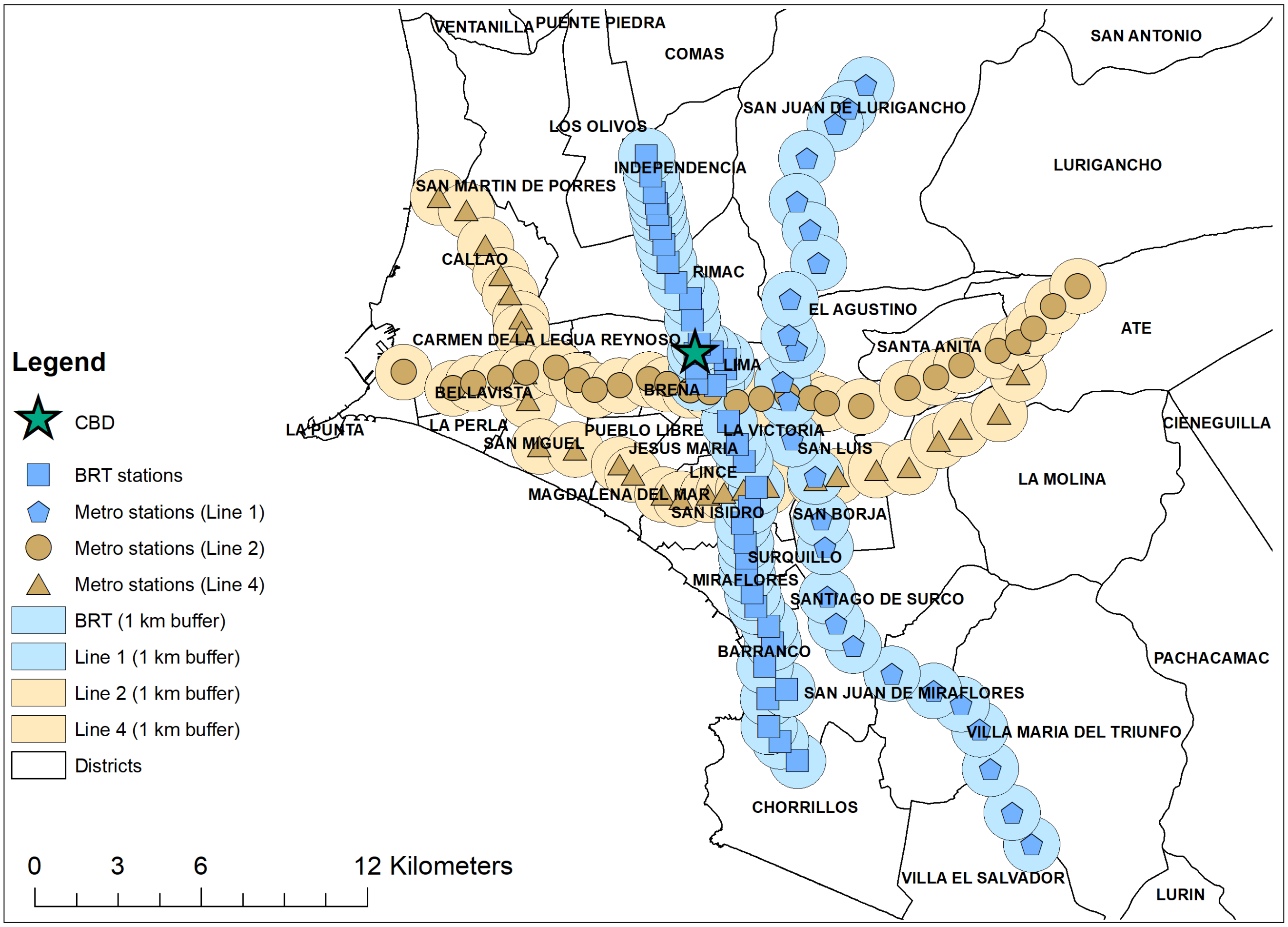
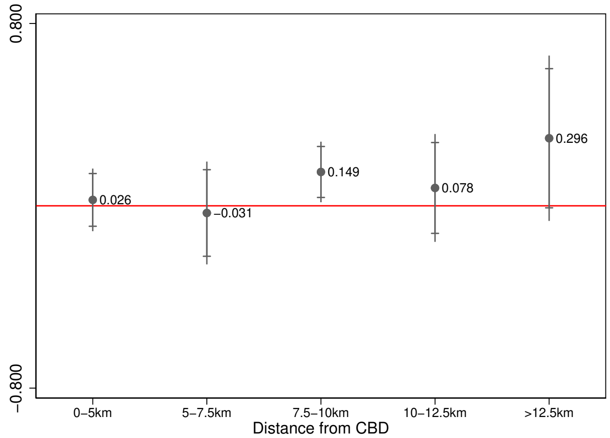


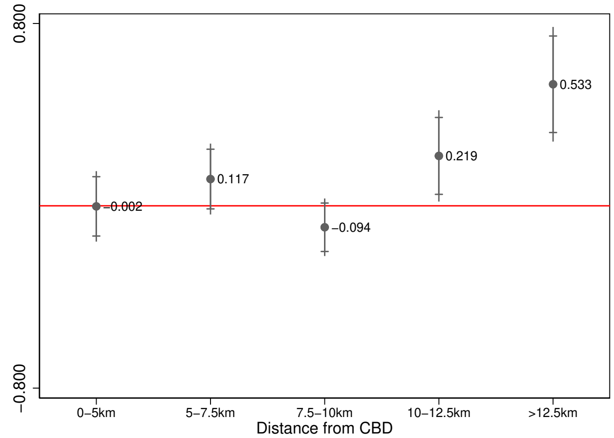
Figure 5: The Reduced Form Impact of Transit Infrastructure on Rents and Expenditure by Distance to the CBD

Panel A: Actual + Hypothetical Rents

Metro

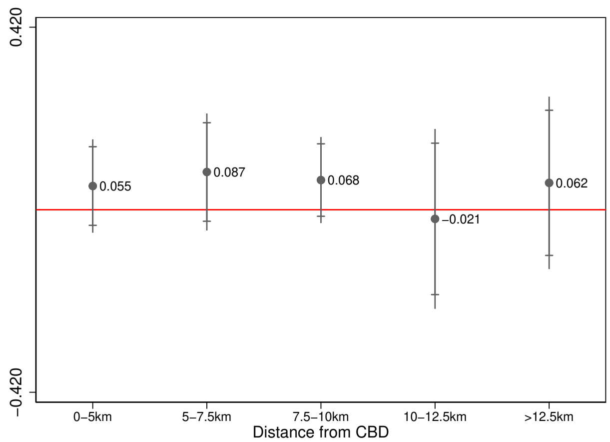


BRT



Panel B: Expenditure

Metro



BRT

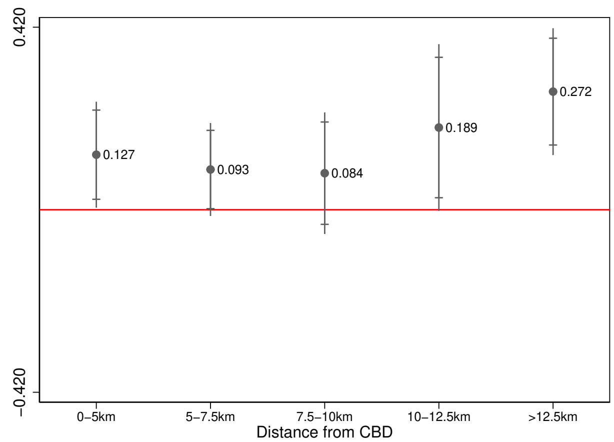


Figure 6: The Reduced Form Impact of Transit Infrastructure on Census Outcomes by Distance to CBD, Part 1

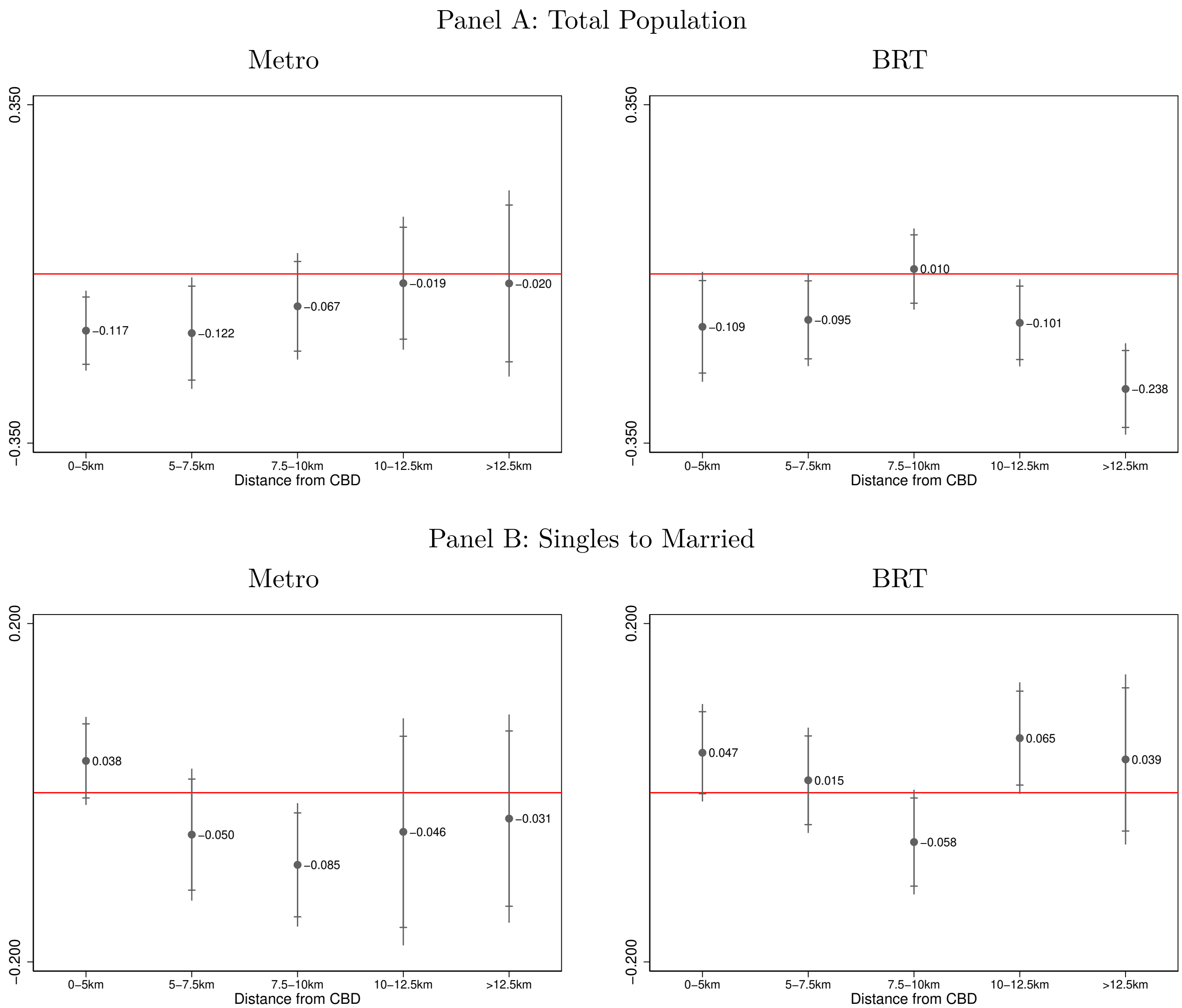
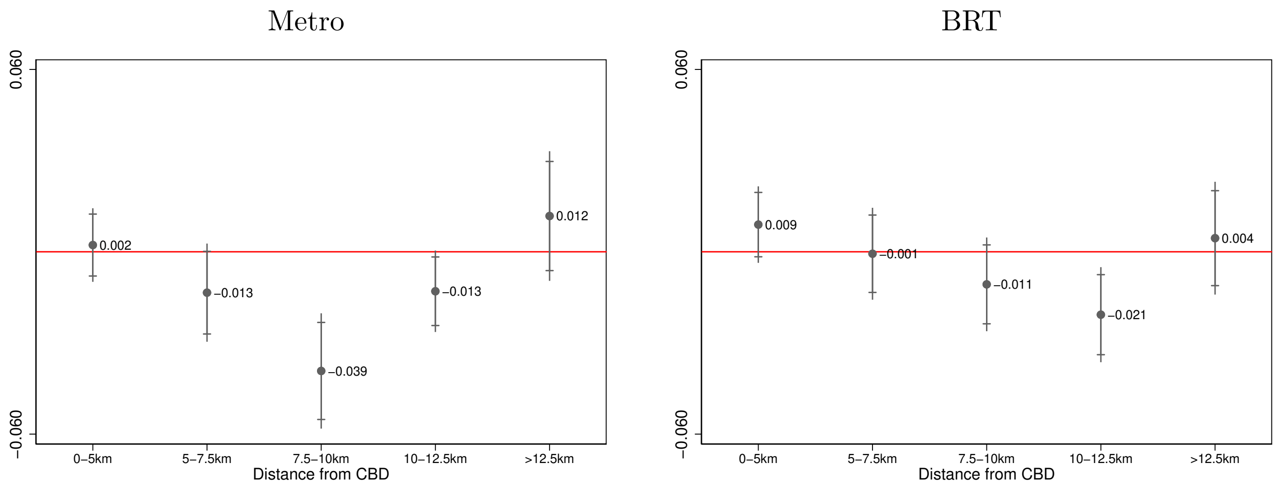


Figure 7: The Reduced Form Impact of Transit Infrastructure on Census Outcomes by Distance to CBD, Part 2

Panel A: Females to Males



Panel B: Dual-Earners to Male Breadwinner

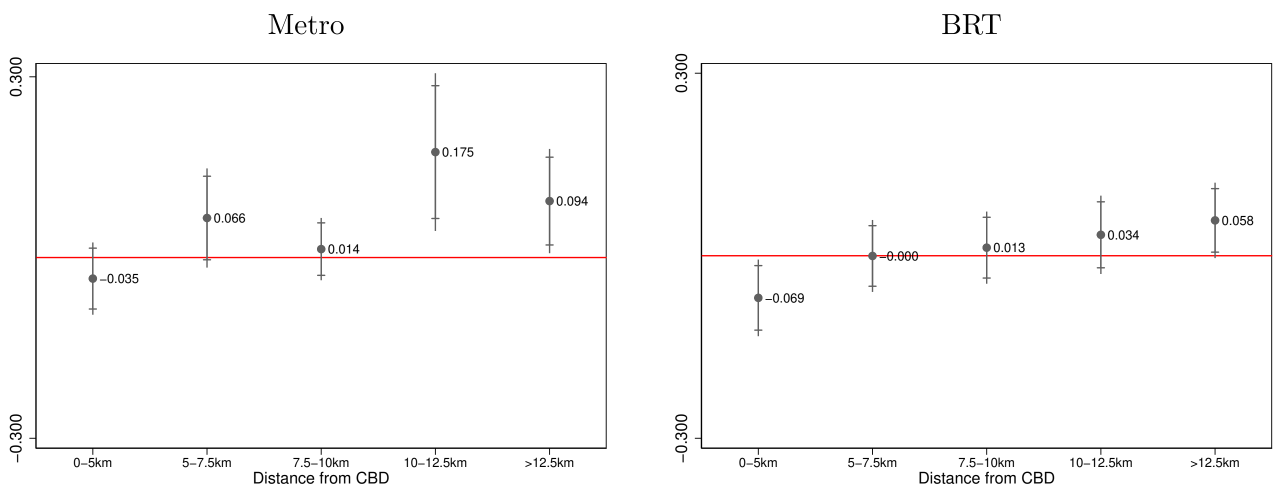
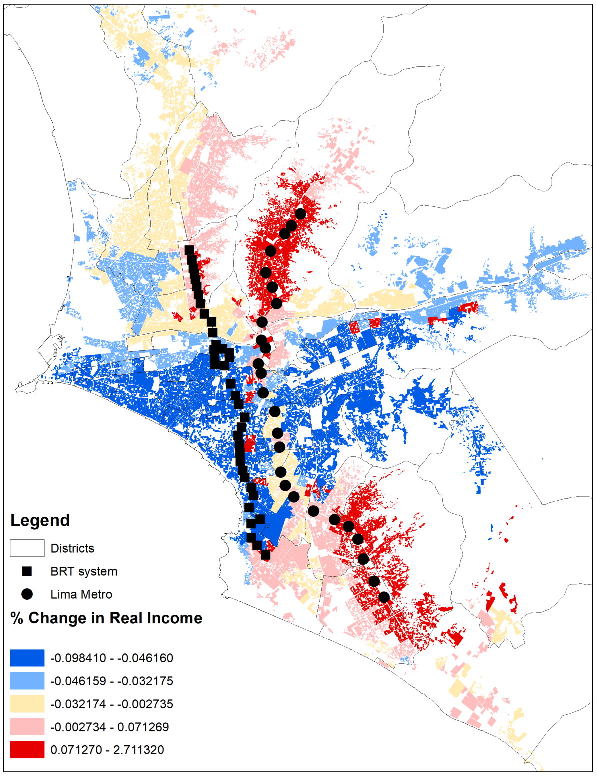


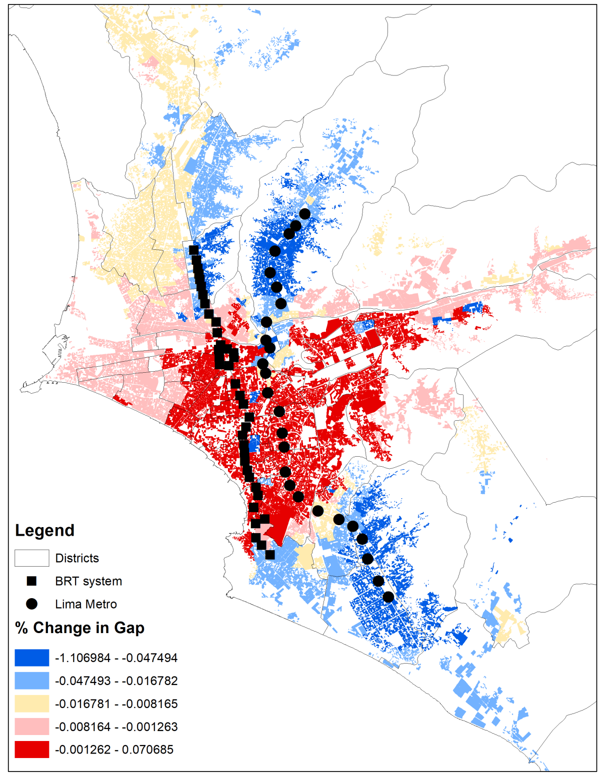
Figure 8: Quantitative Impact of Transit Infrastructure on Earnings and the Earnings Gender Gap with Interdependent Commute Choices

Panel A:

Real Income

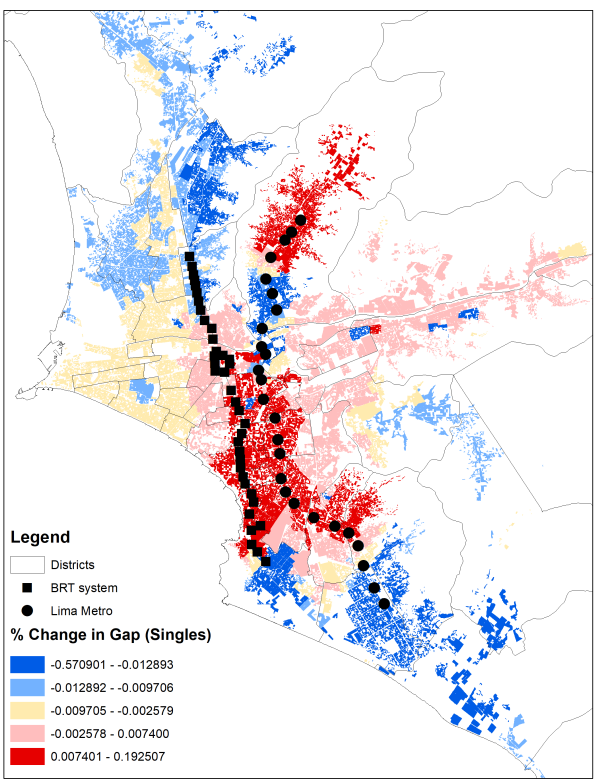


Gender Gap in Real Income



Panel B:

Gender Gap (singles)



Gender Gap (married)

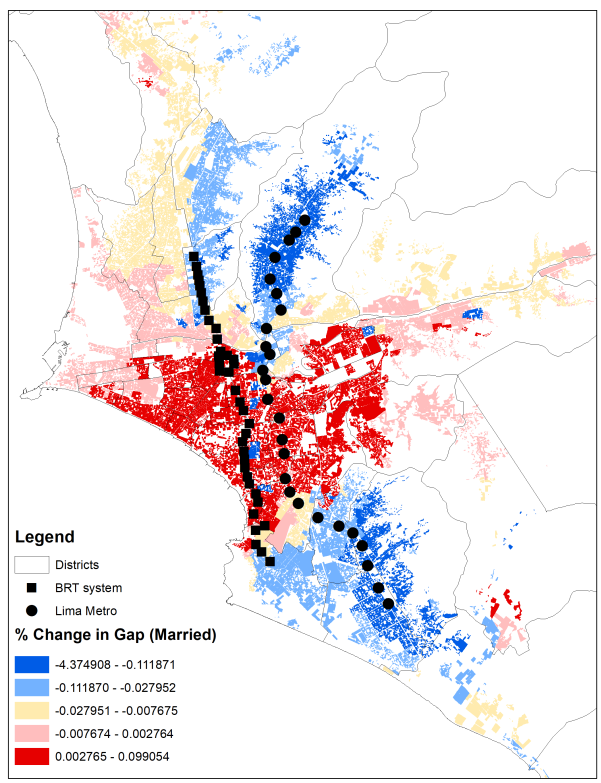
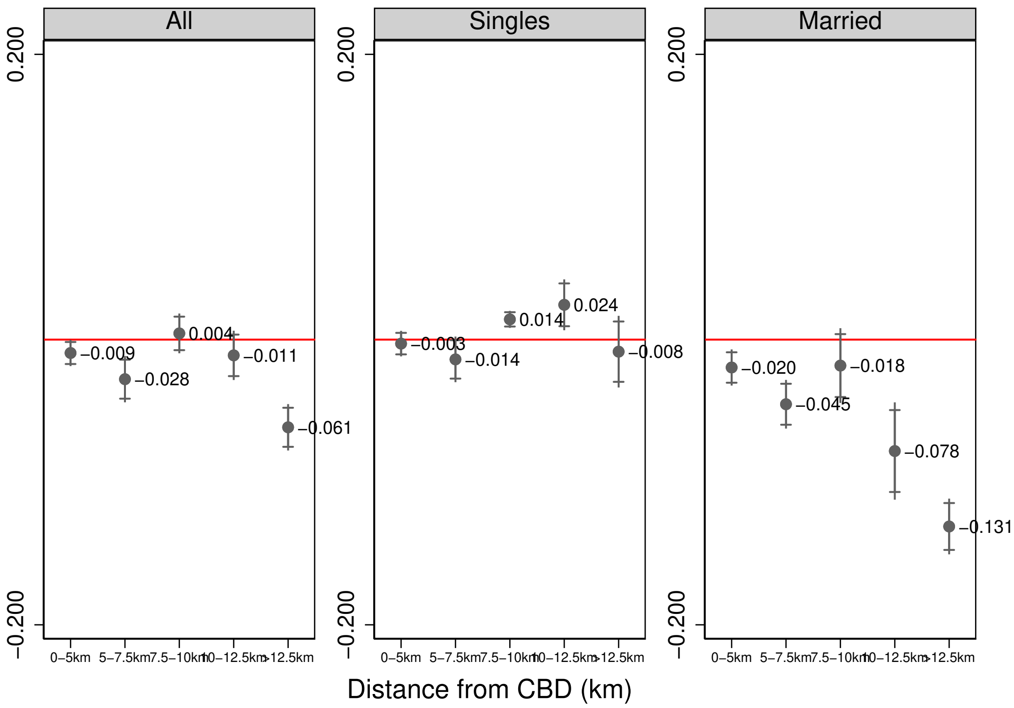


Figure 9: Reduced-form Impact with Interdependent Commute Choices by Distance to the CBD

Panel A: Gap in Aggregate Real Earnings



Panel B: Gap in Real Earnings per Worker

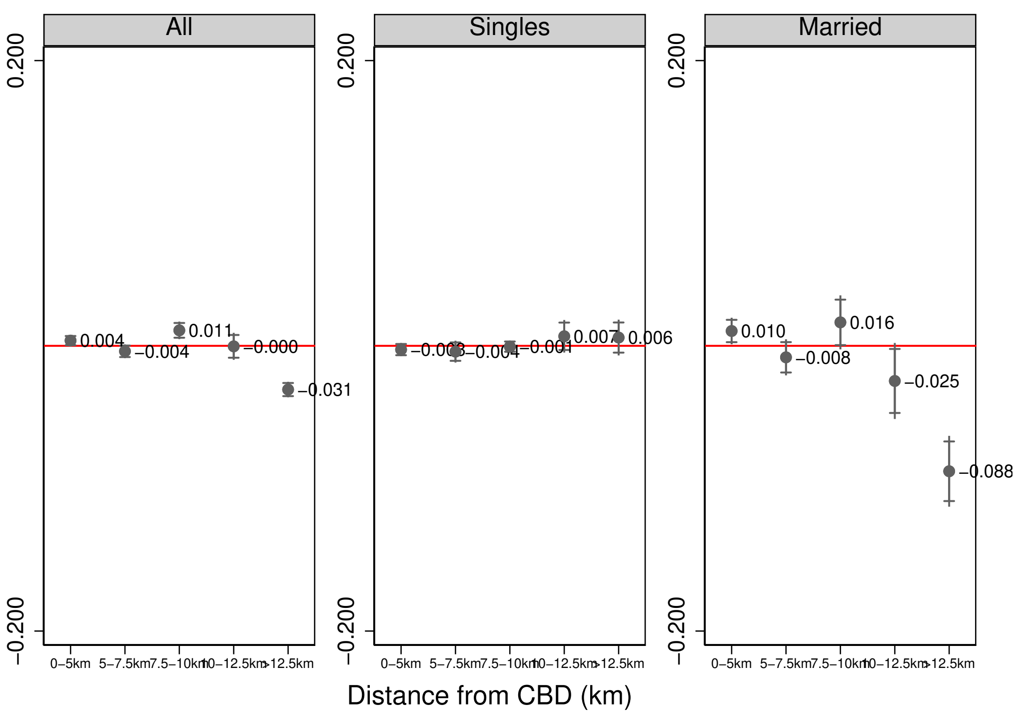
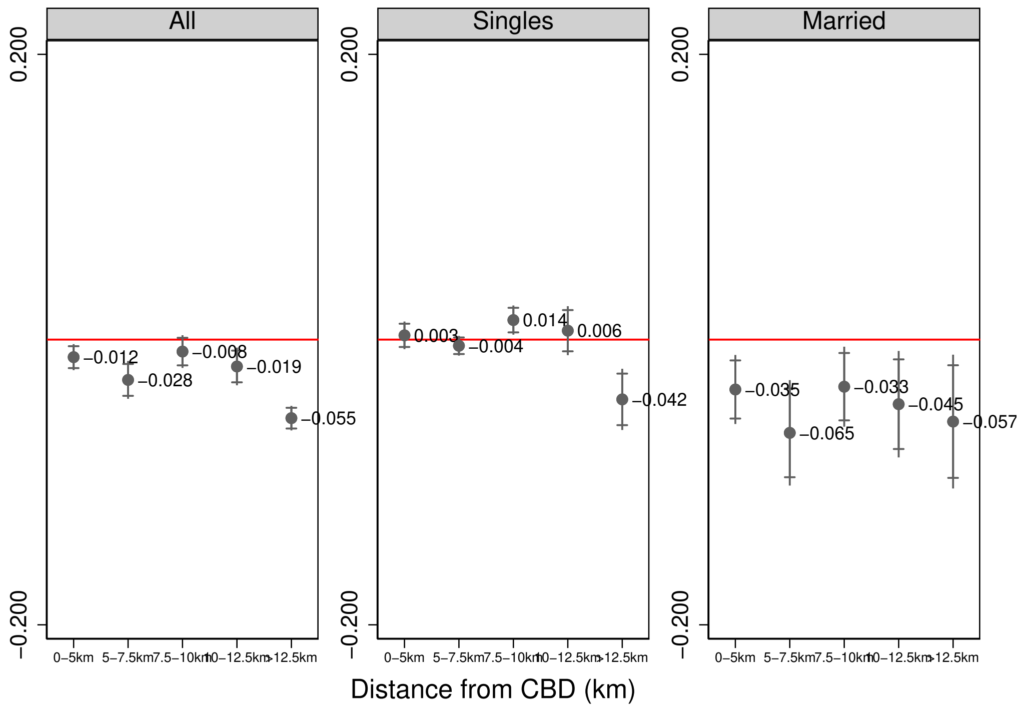


Figure 10: Reduced-form Impact with Independent Commute Choices by Distance to the CBD

Panel A: Gap in Aggregate Real Earnings



Panel B: Gap in Real Earnings per Worker

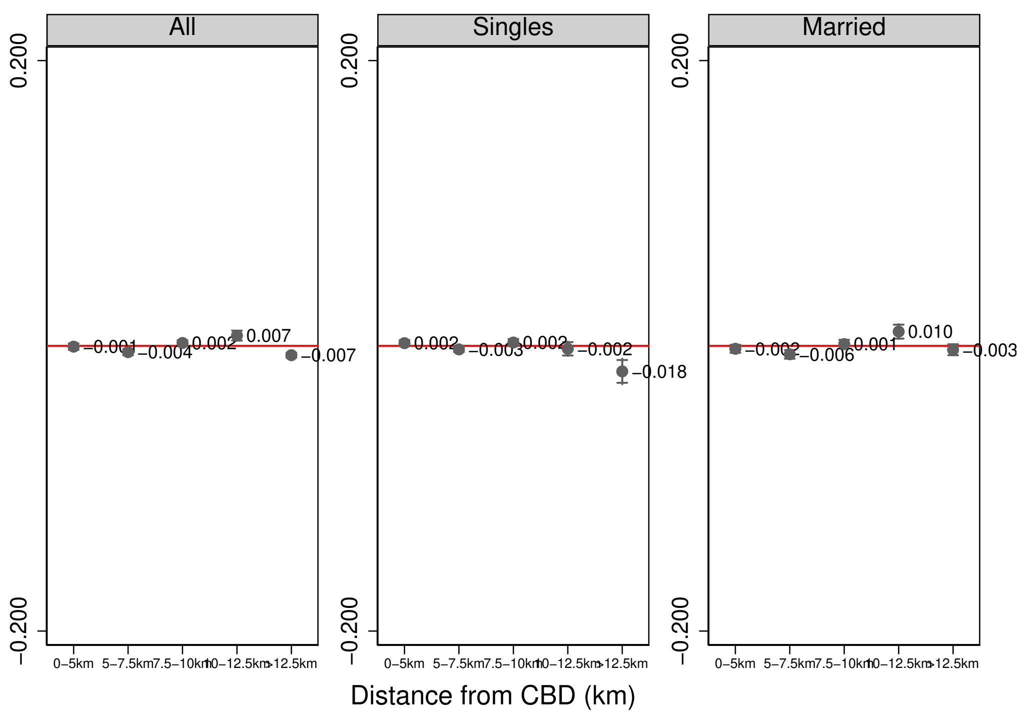
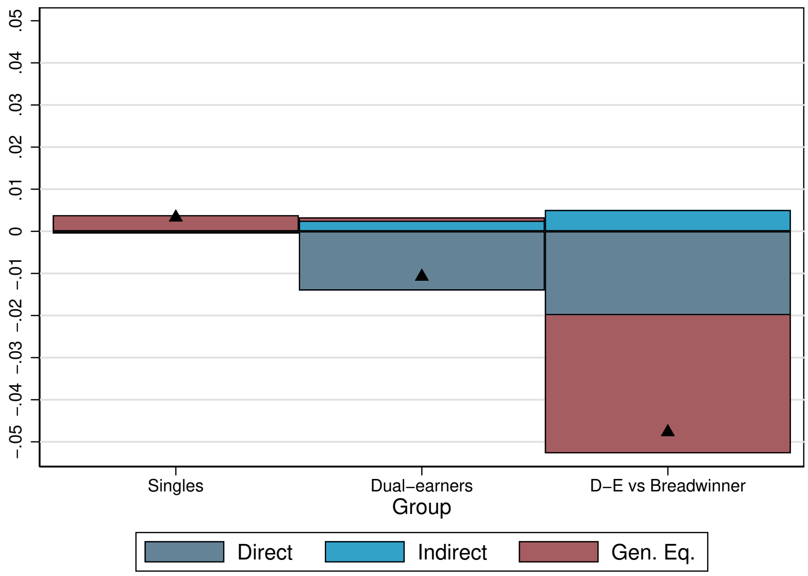
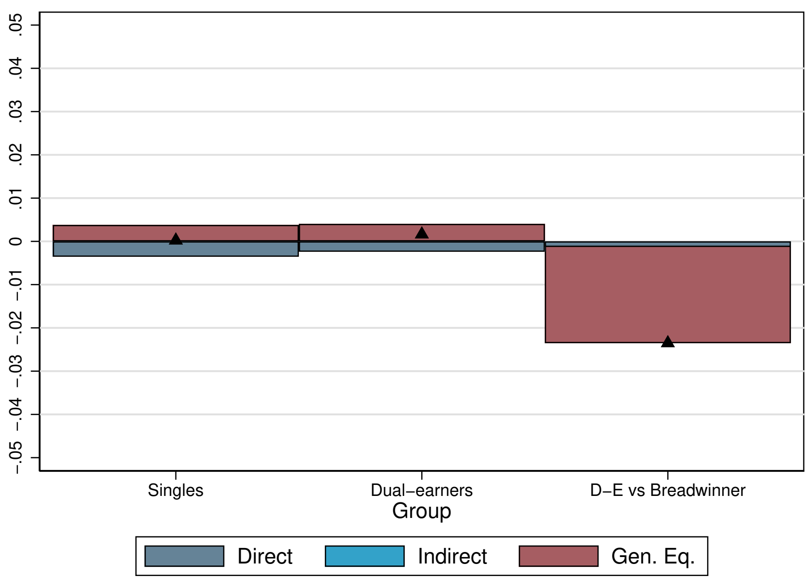


Figure 11: Decomposition of the Aggregate Impacts

Panel A: Interdependent Commute Choices



Panel B: Independent Commute Choices



A1 Data Appendix

In this section, I explain the choices I made when processing the datasets available in this setting.

A1.1 Population and Household Census, 1993, 2007 and 2017

Household Classification. The primary source of population data is the Population and Household Census of 1993, 2007 and 2017. These were conducted by the National Institute of Statistics and Informatics (INEI). Crucially, these data sets contain the population in each block by gender and civil-status. Moreover, the household's roster is observed. Hence, I can observe who lives with whom.

I define a couple as a male and a female adult living in the same two-members-household. In 2017 around 77% of households had two or fewer adult members. In 2007, 75% of households had two or fewer adult members. So, these account for the majority of households. I also include the restriction that at least one of the two should work. Singles are defined in a similar way. I consider households with one adult member. This lone member should work.

For simplicity, I treat individuals living with two or more additional adults as singles (i.e. as if they were roommates). If more than two adults live together, then it is hard to infer who is the couple of whom, and I require this knowledge to compute conditional commute probabilities. In some cases, it can be inferred. For example, when the household head lives with his or her partner. Or if two of the $N > 2$ members report to be married. One can also make some assumptions about the average age gap between the spouses. To avoid further complications, I make the choice of simply treating these individuals as singles. If anything, this choice makes singles to look a bit more alike to couples since I am potentially misidentifying some couples in multi-members households as singles.[51] By considering singles and couples in this simple way I am already improving on what the literature has done. In any case, results are similar if I drop those that report to be married among multi-members households. Finally, my main results consider as adults those between 25 and 65 years old, but results are robust to alternative definitions.

This classification leads to the following distribution of households: (i) single females (31.9% of households in 2017, 29% of households in 2007), (ii) single males (35.9% and 36.1%), (iii) male breadwinner households, that is, households where the wife stays at home while the husband goes to work (13.3% and 17.3%), (iv) female breadwinner households (1.3% and 1.5%), and (v) dual-earner households (17.6% and 16.2%). I calculate the number of households for each of these five groups for each block.

For consistency, I use the same classification across databases.

Commuting Flows. In the 2017 Census, employed individuals were asked to provide the district in which they work. Using this information, along with data on the locations where these individuals reside, I calculate commute probabilities as follows:

$$\pi^g_{j|i} = \frac{N^g_{j|i}}{\sum_k N^g_{k|i}}$$

[51] Less than 50% of individuals in these households report to be married.

where $N^g_{j|i}$ is the number of workers of group g that live in district i and work in district j. For conditional commute probabilities in dual-earner households,

$$\pi^g_{j|il} = \frac{N^g_{j|il}}{\sum_k N^g_{k|il}}$$

where $N^g_{j|il}$ is the number of dual-earner workers of gender g working in j, living i, and whose partners work in l.

A1.2 Economic Census, 2008

The 2008 Economic Census of Peru was conducted by the National Institute of Statistics and Informatics (INEI). This database covers firms of all sizes and all industries, except agriculture and financial services, even those that had not filed their taxes. It focuses exclusively on the firms' operations in 2007. Information and consistency checks were collected by census officers visiting business establishments. To my convenience, the location of the businesses were recorded at a very detailed level, including the specific census block they were located in. I aggregate sectors into seven sectors so that it is easier to match to the Firms' Administrative Data of 2015. In particular, I consider textiles, other manufacture, services, business services, wholesale, retail trade, and transportation services. I aggregate employment counts at the block level for each of the seven industries described above. This way, I get the distribution of economic activity per industry across locations for the first period of the analysis. I scale-up employment by the ratio of employment in the Economic Census to total employment in Lima according to the Population Census in 2007.

A1.3 Firms' Administrative Data, 2015

The Ministry of Production used their raw data and classified more than 1 million firms by industry, geographic location, sales, and number of workers. The database is available at this link. The information relevant to my analysis is the one concerning the industrial classification, the geographic location and the number of workers. Firms were classified into seven categories: 0-5 workers, 6-10 workers, 11-20 workers, 21-50 workers, 51-100 workers, 101-200 workers, and more than 200 workers. To compute the number of workers per firm I simply impute the lower bound of each interval.[52] Then, I aggregate these counts at the block level for each of the following broad industries: textiles, other manufacture, services, business services, wholesale, retail trade, and transportation services. This way, I get the distribution of economic activity per industry across locations for the second period of the analysis. I scale-up employment by the ratio of employment in the Economic Census to total employment in Lima according to the Population Census in 2017.

[52] Results are similar if I instead use the midpoint between the two bounds.

A1.4 Road Network Data and Commute Times

Road network data. I use road network data from Open Street Map. Open Street Map (OSM) is a free, editable map of the world that is created and maintained by a community of volunteer contributors. The project was started in 2004 with the goal of creating a free and open alternative to proprietary mapping services, such as Google Maps. The data in OSM is collected from a variety of sources, including GPS tracks, aerial imagery, and manual surveying. The data is then added to the OSM database and is available for anyone to use. The quality of OpenStreetMap data can vary depending on the region and the level of community engagement. In general, OSM data tends to be more accurate and up-to-date in areas with a high level of community engagement, such as densely populated urban areas and regions with a strong OSM community. In these areas, the map is regularly updated with new data and errors are quickly corrected. Importantly, the OSM data contains information on road type or classification, which I use to impute road speeds.

Commute times. I use the data on road type to input speeds during rush hour according to the following classification: motorway (50 km/h), primary (30 km/h), residential (15 km/h), secondary (20 km/h), and tertiary (10 km/h). This classification is based on a report for the Ministry of Transport and Communications done by the Japan International Cooperation Agency in 2013 (JICA, 2013). Additionally, I manually input speeds in certain roads, the BRT, and the Metro following the same report (see Table 8). Then, I run ArcGis's Network Analyst tool to compute the optimal route from origins to destination to generate an origin-destination matrix at the block level. I utilize this tool to compute commute times in absence of the Metro and the BRT and when the two are functional. Sometimes, I aggregate these times to the district level by taking the median value across zones within districts. I show that these times are well correlated with times reported in commuting surveys (see Figure B.3).

The matrix of commute times used in the analysis is an approximation of actual commute times. Ideally, the routes of public buses and road speeds would be used to calculate how long it takes to travel between locations using public transportation. However, data on bus routes is not readily available in this context. Despite this limitation, the commute elasticities, which link commute times to costs, are still estimated using the approximation of commute times and probabilities. These elasticities already take into account any potential differences between the estimated and actual commute times. For example, if actual commute times are twice as long as the estimated times, the estimated elasticities would be half as large. The only concern would be if the error in the approximation of commute times is correlated with a blocks' location or some characteristic, which could lead to problems.

Table 8: Manually imputed speeds

Road	Speed
Avenida Brasil	15 km/h
Avenida Javier Prado Oeste	15 km/h
Avenida La Marina	15 km/h
Avenida Faustino Sanchez Carrión	15 km/h
Avenida Almirante Miguel Grau	15 km/h
Avenida Abancay	15 km/h
Avenida Arequipa	15 km/h
Avenida de Tomás Marsano	15 km/h
Avenida República de Panamá	15 km/h
Avenida Arica	15 km/h
Avenida Aviación	15 km/h
Avenida República de Venezuela	15 km/h
Avenida Bolognesi	15 km/h
Avenida Prolongación Paseo de la República	15 km/h
Avenida Defensores del Morro	15 km/h
Avenida Universitaria	15 km/h
Metro Lines	45 km/h
Bus Rapid Transit System	35 km/h

A1.5 National Household Survey, 2007-2017

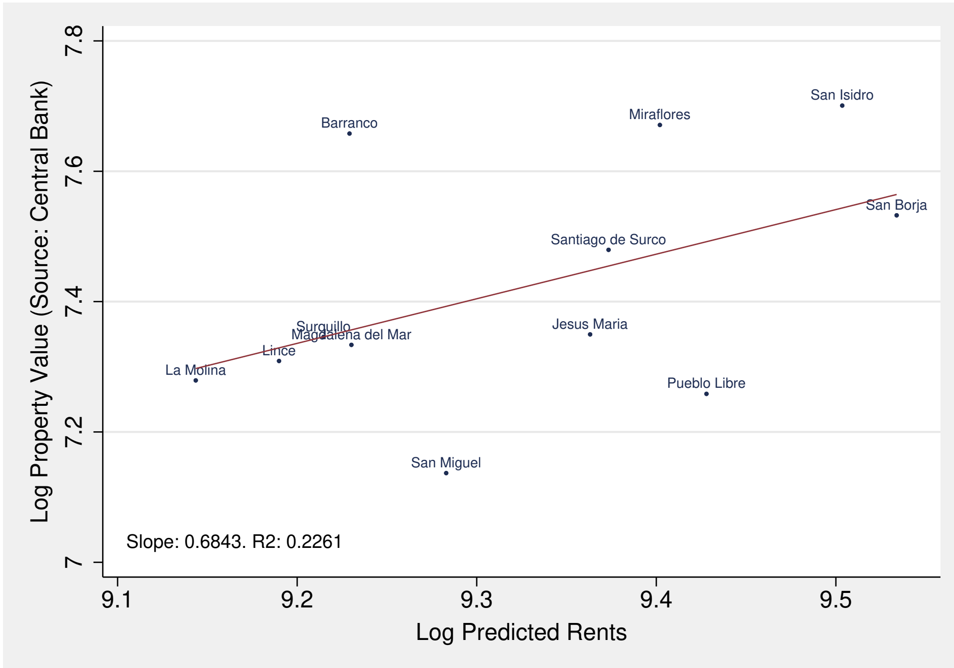
The National Institute of Statistics and Informatics (INEI) in Peru conducts the Peruvian National Household Survey (ENAHO) annually. The sample size of the survey varies each year, with a range of 26,000 to 40,000 households. The survey provides a wealth of information on individual income and demographics, including details on primary and secondary activities such as wages and hours worked. One of the survey's key benefits is that households are geocoded at the cluster level, with each cluster containing about 140 households. This level of granularity allows for accurate assignment of households to census blocks. Additionally, the survey collects valuable information from household heads, such as the estimated rental value of their home, total household income, and expenditures.

I use this data to predict how property values vary with centrality, and other block and dwelling characteristics.[53] Then, I use the estimated coefficients to project property values into census' blocks. I assign data on years 2007-2010 to the first period considered in the analysis, and data on years 2014-2017 to the second period. In Figure 12, I compare the log of the average value of predicted rents by district to data on the log of property values compiled by the Central Bank of Peru for a subset of districts. These averages are computed over years 2014-2017. Figure 12 shows that the two measures are well correlated. Finally, since geocoded data on rents by firms is not available in this setting, I assume that there are no differential rents due to land regulations, and so, the returns to floorspace

[53] I achieve and R^2 of around 60% across the two periods considered in the analysis.

use in residential and commercial activities are the same, conditional on a block's location.

Figure 12: Property Values versus Predicted Rents by District



A1.6 Land Use Data, 2020

The Metropolitan Planning Institute collects actual land use data, but it is only available for the year 2020 and not for previous years. As a result, when required, I am assuming that land use in 2007 is the same as it was in 2017.[54] Despite this assumption, it does not have any impact on the estimation of any parameters in the model. This is because this data is only used to recover unobservables on the housing side, which is modeled in a straightforward manner following Tsivanidis (2021). These unobsevables are not an input in the estimation of any of the parameters. Additionally, the counterfactual analysis is not affected as I am using 2017 data as a baseline when removing the Metro and BRT from Lima. Therefore, the lack of land use data for 2007 does not affect this exercise. The lack of this data would be of significance if I were to perform a counterfactual analysis using the year 2007 as baseline.

[54] An alternative option would be to use data on floor space zoning, which is available for earlier years.

A2 Solution Algorithm

Corner solutions and observed prices

Given that I have agglomeration forces in the model, it is possible to have corner allocations for some locations. Without these externalities it is only possible if the exogenous component of productivity or of amenities is equal to zero. With externalities, even if the exogenous component of productivity or of amenities is positive, I can have corner solutions as the endogenous part, which enters multiplicatively, can be equal to zero when either employment or residence are equal to zero.

This implies that the groups $\mathcal{I}_{FR}$, $\mathcal{I}_F$, and $\mathcal{I}_R$ are endogenously determined:

$$\begin{aligned}
\mathcal{I}_{FR} &= \{i : u_{ik} > 0, A_{is} > 0 \text{ for some } k \in mf, m, f \text{ and for some } s\} \\
\mathcal{I}_R &= \{i : u_{ik} > 0, A_{is} = 0 \text{ for some } k \in mf, m, f \text{ and for all } s\} \\
\mathcal{I}_F &= \{i : u_{ik} = 0, A_{is} > 0 \text{ for all } k \in mf, m, f \text{ and for some } s\}
\end{aligned}$$

Following Ahlfeldt et al. (2015), we can define the observed price of floor space ($\mathbb{r}_i$). We can summarize the relationship between the observed price of floor space ($\mathbb{r}_i$), the price of commercial floor space (r_{F_i}), the price of residential floor space (r_{R_i}), and land use as (ϑ_i):

$$\mathbb{r}_i = \begin{cases} \zeta_{F_i} r_{F_i} & \zeta_{F_i} = 1, \quad i \in \mathcal{I}_F = \{i : u_{ik} = 0 \text{ for all k}, A_{is} > 0 \text{ for some s}\} \\ \zeta_{F_i} r_{F_i} & \zeta_{F_i} = 1 - \tau_i, \quad i \in \mathcal{I}_{FR} = \{i : u_{ik} > 0 \text{ for some k}, A_{is} > 0 \text{ for some s}\} \end{cases} \tag{21}$$

$$\mathbb{r}_i = \begin{cases} \zeta_{R_i} r_{R_i} & \zeta_{R_i} = 1, \quad i \in \mathcal{I}_R = \{i : u_{ik} > 0 \text{ for some k}, A_{is} = 0 \text{ for all s}\} \\ \zeta_{R_i} r_{R_i} & \zeta_{R_i} = 1, \quad i \in \mathcal{I}_{FR} = \{i : u_{ik} > 0 \text{ for some k}, A_{is} > 0 \text{ for some s}\} \end{cases}$$

where ζ_{F_i} and ζ_{R_i} relate observed floor prices to commercial and residential floor prices. $\mathcal{I}_F$ is the set of locations specialized in commercial activity ($1 - \vartheta_i = 1$), $\mathcal{I}_R$ is the set of locations specialized in residential activity ($\vartheta_i = 1$), and $\mathcal{I}_{FR}$ is the set of locations with both commercial and residential activity ($\vartheta_i \in (0, 1)$). Note that according to equation 21, the relationship between he observed price of floor space ($\mathbb{r}_i$), the price of commercial floor space (r_{F_i}), the price of residential floor space (r_{R_i}), and land use as (ϑ_i), are a function only of the exogenous locational characteristics given by the vector defined by A_{is}, u_{ig}, and τ_i.

Algorithm

The system of equations defined above can be solved using the following algorithm.

1. Guess vector $w_g^0, r^0, \vartheta^0, u_{mf}^0, u_m^0, u_f^0, A_s^0, \mu^0$.
   - Vector of prices should have normalization, e.g. $w_{1m} = 1$, so divide the vector of wages and rents by w_{1m}.
2. Given a vector $w_g^t, r^t, \vartheta^t, u_{mf}^t, u_m^t, u_f^t, A_s^t, \mu^t$:

(a) Given our guess for productivity and amenities, we can determine the group indicator variables:

$$\begin{aligned}
\mathcal{I}_{FR} &= \{i : u_{ik} > 0, A_{is} > 0 \text{ for some } k \in mf, m, f \text{ and for some } s\} \\
\mathcal{I}_{R} &= \{i : u_{ik} > 0, A_{is} = 0 \text{ for some } k \in mf, m, f \text{ and for all } s\} \\
\mathcal{I}_{F} =& \quad \{i : u_{ik} = 0, A_{is} > 0 \text{ for all } k \in mf, m, f \text{ and for some } s\}
\end{aligned}$$

(b) Construct rents for this iteration:

$$r_{R_i} = \begin{cases} r_i \times (1 - \tau_i) & \text{if} i \in \mathcal{I}_{FR} \\ r_i & \text{if} i \in \mathcal{I}_{R} \\ . & \text{if} i \in \mathcal{I}_{F} \end{cases} \quad r_{F_i} = \begin{cases} r_i & \text{if} i \in \mathcal{I}_{FR} \\ . & \text{if} i \in \mathcal{I}_{R} \\ r_i & \text{if} i \in \mathcal{I}_{F} \end{cases}$$

(c) Compute total rents:

$$E = \sum_i r_{R_i} H_{R_i} + r_{F_i} H_{F_i}$$

(d) Compute residence market access for single households:

$$\Psi_{R_i^g} = \sum_{l'} \left(\frac{d^g_{il'}}{w_{l'g}} \right)^{-\theta^g}, \forall g \in \{m, f\}$$

(e) Compute residence market access for couples:

i. In male breadwinner households:

$$\Phi^h_{R_{i,\ell=0}} = \sum_{l'} \left(\frac{d^h_{il}}{w_{lm}} \right)^{-\theta^h}$$

ii. In dual-earner households:

$$\begin{aligned}
\Phi^h_{R_{i,\ell=1,j'}} &= \sum_l \left(\frac{d^h_{il}}{w_{lm} + w_{j'f}} \right)^{-\theta^h} \\
\Phi^w_{R_{i,\ell=1}} &= \sum_l \left(\frac{\left(\Phi^h_{R_{i,\ell=1,l}}\right)^{1/\theta^h}}{d^w_{il}} \right)^{\theta^w}
\end{aligned}$$

iii. On average:

$$\Psi_{R_i^{mf}} = \left[\left(\Phi_{R^w_{i,\ell=1}}\right)^{\frac{\nu}{\theta^w}} + \left(\xi^{mf}_{i\ell=0}\right)^{\nu} \left(\Phi_{R^h_{i,\ell=0}}\right)^{\frac{\nu}{\theta^h}} \right]^{1/\nu}$$

(f) Compute supply of residents to each location for each gender-by-household type:

$$
\begin{aligned}
N_{R_i^{k,k}} &= \pi_i^{k,k} N^{k,k}, k \in \{m, f\} \\
N_{R_i^{mf}} &= \underbrace{(1-\mu_i)\,\pi_i^{mf} N^{mf}}_{N_{R_{i,0}^{mf}}} + \underbrace{(\mu_i)\,\pi_i^{mf} N^{mf}}_{N_{R_{i,1}^{mf}}}
\end{aligned}
$$

where

$$
\pi_i^g = \frac{\left(\frac{r_{R_i}^{1-\beta^g}}{u_i^g \Psi_{R_i^g}^{1/\theta g}}\right)^{-\eta^g}}{\sum_l \left(\frac{r_{R_l}^{1-\beta^g}}{u_l^g \Psi_{R_l^g}^{1/\theta g}}\right)^{-\eta^g}}
$$

and

$$
\pi_i^{mf} = \frac{\left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_i^{mf} \Psi_{R_i^{mf}}^{1/\nu}}\right)^{-\eta^{mf}}}{\sum_l \left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_l^{mf} \Psi_{R_l^{mf}}^{1/\nu}}\right)^{-\eta^{mf}}}
$$

(g) Compute commuting probabilities:

i. For singles:

$$
\pi_{j|i}^g = \frac{\left(\frac{d_{ij}^g}{w_{jg}}\right)^{-\theta, g}}{\Phi_{R_i^g}}
$$

ii. For males in male-breadwinner households:

$$
\pi_{j|i,\ell=0}^h = \frac{\left(\frac{d_{ij}^h}{w_{jm}}\right)^{-\theta^h}}{\Phi_{R_{i,\ell=0}}^h}
$$

iii. For females in dual-earner households:

$$
\pi_{j'|i,\ell=1}^w = \frac{\left(\frac{d_{ij'}^w}{\left(\Phi_{R_{i,\ell=1,j'}}^h\right)^{1/\theta^h}}\right)^{-\theta^w}}{\Phi_{R_{i,\ell=1}}^w}
$$

iv. For males in dual-earner households:

$$\pi^h_{j|i,\ell=1} = \sum_l \pi^h_{j|i,\ell=1,l} \cdot \pi^w_{l|i,\ell=1}$$

where

$$\pi^h_{j|i,\ell=1,j'} = \frac{\left(\frac{d^h_{ij}}{w_{jm}+w_{j'f}}\right)^{-\theta^h}}{\Phi^h_{R_{i,\ell=1,j'}}}$$

(h) Compute labor supply:

$$\begin{aligned}
L_{F_{jm}} &= \sum_i \left\{\pi^{m,m}_{j|i} N_{R^{m,m}_i}\right\} + \sum_i \left\{\pi^h_{j|i,\ell=0} N_{R^{mf}_{i,0}}\right\} + \sum_i \left\{\pi^h_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} \\
L_{F_{jf}} &= \sum_i \left\{\pi^{f,f}_{j|i} N_{R^{f,f}_i}\right\} + \sum_i \left\{\pi^w_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\}
\end{aligned}$$

(i) Compute total income:

$$\begin{aligned}
y^{mf}_{i,\ell=1} &= \underbrace{\sum_j \left(w_{j,m}\pi^h_{j|i,\ell=1}\right) N_{R^{mf}_{i,1}}}_{y^h_{i,\ell=1}} + \underbrace{\sum_j \left(w_{j,f}\pi^w_{j|i,\ell=1}\right) N_{R^{mf}_{i,1}}}_{y^w_{i,\ell=1}} \\
y^{mf}_{i,\ell=0} &= \sum_j \left(w_{j,m}\pi^h_{j|i,\ell=0}\right) N_{R^{mf}_{i,0}} \\
y^g_i &= \sum_j \left(w_{j,g}\pi^g_{j|i}\right) N_{R^g_i}
\end{aligned}$$

(j) Compute price index:

$$\begin{aligned}
p_{js} &= W^{\alpha_s}_{js} r^{1-\alpha_s}_{Fj} / A_{js} \\
P &= \left(\sum \left(W^{\alpha_s}_{js} r^{1-\alpha_s}_{Fj} / A_{js}\right)^{1-\sigma_D}\right)^{\frac{1}{1-\sigma_D}} \\
W_{js} &= \left(\sum_g \alpha^\sigma_{sg} w^{1-\sigma}_{jg}\right)^{\frac{1}{1-\sigma}}
\end{aligned}$$

(k) Update guesses using demand side:

i. Update labor force participation:

$$\mu^{t+1}_{i,mf} = \frac{\left[\left(\Phi^w_{R_{i,\ell=1}}\right)^{-1/\theta^w}\right]^{-\nu}}{\Psi_{R^{mf}_i}}$$

ii. Update endogenous productivities:

$$A_{js}^{t+1} = \bar{A}_{js}\left(L_{F_{js}}\right)^{\phi_{Fs}}$$

iii. Plugging labor supply into labor demand to update new wages:

$$\tilde{w}_{jg} = \left(\frac{\sum_s \left(\frac{1}{\alpha_{sg} W_{js}}\right)^{-\sigma} M_{js}}{L_{F_{jg}}}\right)^{1/\sigma}$$

where:

$$\begin{aligned}
M_{js} &= \alpha_s \frac{X_{js}}{W_{js}} \\
X_{js} &= \frac{p_{js}^{1-\sigma_D}}{P^{1-\sigma_D}} X \\
X = PC &= \sum_i PC_i + \sum_i \left(H_{R_i} r_{R_i} + H_{F_i} r_{F_i}\right) \\
&= \beta^m y_i^{m,m} + \beta^f y_i^{f,f} + \beta^{mf} y_{i,\ell=1}^{mf} + \beta^{mf} y_{i,\ell=0}^{mf} + E
\end{aligned}$$

iv. Update observed rents. To do so, note that aggregate expenditure on housing in location i by households is:

$$\begin{aligned}
E_{Ri} &= r_{R_i} H_{R_i} \\
&= \left(1-\beta^m\right) y_i^{m,m} + \left(1-\beta^f\right) y_i^{f,f} + \left(1-\beta^{mf}\right) y_{i,\ell=1}^{mf} + \left(1-\beta^{mf}\right) y_{i,\ell=0}^{mf}
\end{aligned}$$

Similarly, note that aggregate spending on floorspace by firms is:

$$H_{F_i} = \sum_s \left(1-\alpha_s\right) \frac{X_{js}}{r_{F_j}}$$

Therefore, our new guess for observed rents would be as follows:

$$\begin{aligned}
\text{if i} \in \mathcal{I}_R, \text{then} r_i^{t+1} &= E_{Ri}/H_i \\
\text{if i} \in \mathcal{I}_F, \text{then} r_i^{t+1} &= X_i/H_i \\
\text{if i} \in \mathcal{I}_{FR}, \text{then} r_i^{t+1} &= \frac{\frac{E_{Ri}}{(1-\tau_i)} + X_i}{H_i}
\end{aligned}$$

v. Update land allocation:

$$\begin{aligned}
\text{if i} \in \mathcal{I}_R, \text{then} \vartheta_i &= 1 \\
\text{if i} \in \mathcal{I}_F, \text{then} \vartheta_i &= 0 \\
\text{if i} \in \mathcal{I}_{FR}, \text{then} \vartheta_i &= \frac{\frac{E_{Ri}}{(1-\tau_i)}}{\frac{E_{Ri}}{(1-\tau_i)} + X_i}
\end{aligned}$$

(l) Normalize new vector of prices.

(m) If $|| \left(w_g^{t+1}, r^{t+1}, \vartheta^{t+1}, u_{mf}^{t+1}, u_m^{t+1}, u_f^{t+1}, A_s^{t+1}, \mu^{t+1}\right) - \left(w_g^t, r^t, \vartheta^t, u_{mf}^t, u_m^t, u_f^t, A_s^t, \mu^t\right) || \leq tolerance$, then stop. Otherwise, set:

$$\begin{aligned}\left(w_g^{t+1}, r^{t+1}, \vartheta^{t+1}, u_{mf}^{t+1}, u_m^{t+1}, u_f^{t+1}, A_s^{t+1}, \mu^{t+1}\right) &= \rho \left(w_g^{t+1}, r^{t+1}, \vartheta^{t+1}, u_{mf}^{t+1}, u_m^{t+1}, u_f^{t+1}, A_s^{t+1}, \mu^{t+1}\right) \\ &+ (1-\rho)\left(w_g^t, r^t, \vartheta^t, u_{mf}^t, u_m^t, u_f^t, A_s^t, \mu^t\right)\end{aligned}$$

where $\rho \in (0,1)$.

(n) Return to step 2.a

A3 Proofs

A3.1 Proposition 1

PROPOSITION: *Given data on residence by gender and household type, $N_{R^k_{i,\ell}}$, employment by gender $L_{F_{jg}}$ and commute costs d^k_{ij}, and parameters θ^k, there exists a vector of wages w_{jg} up to scale that rationalizes the observed data as a model equilibrium. Additionally, if one does not observe $L_{F_{jg}}$ but rather employment by industry $L_{F_{js}}$ and parameters α_{sg} and σ, there exists a vector of wages w_{jg} up to scale that rationalizes the observed data as a model equilibrium.*

Proof. Recall that labor supply is:

$$
\begin{aligned}
L_{F_{jm}} &= \sum_i \left\{ \pi^{m,m}_{j|i} N_{R^{m,m}_i} \right\} + \sum_i \left\{ \pi^h_{j|i,\ell=0} N_{R^{mf}_{i,0}} \right\} + \sum_i \left\{ \pi^h_{j|i,\ell=1} N_{R^{mf}_{i,1}} \right\} \\
L_{F_{jf}} &= \sum_i \left\{ \pi^{f,f}_{j|i} N_{R^{f,f}_i} \right\} + \sum_i \left\{ \pi^w_{j|i,\ell=1} N_{R^{mf}_{i,1}} \right\}
\end{aligned}
$$

Or

$$
\begin{aligned}
L_{F_{jm}} &= w^{\theta^m}_{jm} \sum_i \left[\frac{N_{R^m_i} \left(d^m_{ij}\right)^{-\theta^m}}{\sum_{l'} w^{\theta^m}_{jm} \left(d^m_{il'}\right)^{-\theta^m}} \right] + w^{\theta^h}_{jm} \sum_i \left[\frac{N_{R^{mf}_{i,0}} \left(d^h_{ij}\right)^{-\theta^h}}{\sum_l w^{\theta^h}_{lm} \left(d^h_{il}\right)^{-\theta^h}} \right] \\
&+ \sum_i \left[\frac{N_{R^{mf}_{i,1}} \left(d^h_{ij}\right)^{-\theta^h}}{\sum_{l'} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{l'f}} \right)^{-\theta^h} \right)^{\theta^w/\theta^h} \left(d^w_{il'}\right)^{-\theta^w}} \times \right. \\
&\times \left. \sum_l \frac{\left(d^w_{il}\right)^{-\theta^w} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{lf}} \right)^{-\theta^h} \right)^{\theta^w/\theta^h - 1}}{\left(w_{jm} + w_{lf}\right)^{-\theta^h}} \right]
\end{aligned}
$$

and

$$
\begin{aligned}
L_{F_{jf}} &= w^{\theta^f}_{jf} \sum_i \left[\frac{N_{R^f_i} \left(d^f_{ij}\right)^{-\theta^f}}{\sum_{l'} w^{\theta^f}_{jf} \left(d^f_{il'}\right)^{-\theta^f}} \right] \\
&+ \sum_i \left[\frac{N_{R^{mf}_{i,1}} \left(d^w_{ij}\right)^{-\theta^w}}{\sum_{l'} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{l'f}} \right)^{-\theta^h} \right)^{\theta^w/\theta^h} \left(d^w_{il'}\right)^{-\theta^w}} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm} + w_{jf}} \right)^{-\theta^h} \right)^{\theta^w/\theta^h} \right]
\end{aligned}
$$

Inuitively, if one could observe residence supply by gender and household type, $N^{\ell}_{R_{i,k}}$, and employment by gender, $L_{F_{jg}}$, then one could solve the system defined above.

Additionally, if one does not observe $L_{F_{jg}}$ but rather $L_{F_{js}}$ and knows parameters α_{sg} and σ, from

FOCs of the firm, one gets:

$$L_{F_{jg}} = \sum_s \frac{\left(w_{jg}/\alpha_{sg}\right)^{-\sigma}}{\sum_{h\in m,f}\left(w_{jh}/\alpha_{sh}\right)^{-\sigma}} L_{F_{js}}$$

We can plug this expression to the equations above, and solve the system. In what is next, I follow Mas-Collel, Whinston, and Green (1995) and Allen, and Arkolakis (2015) to prove existence. While existence is assured, uniqueness is not unless additional conditions are imposed. Let $\mathcal{Z}_{jm}$ and $\mathcal{Z}_{jf}$ be defined as:

$$\begin{aligned}
\mathcal{Z}_{jm} &= \sum_s \frac{\left(w_{jm}/\alpha_{sm}\right)^{-\sigma}}{\sum_{g\in m,f}\left(w_{jg}/\alpha_{sg}\right)^{-\sigma}} L_{F_{js}} - \sum_i \left\{\pi^m_{j|i} N_{R^m_i}\right\} + \sum_i \left\{\pi^h_{j|i,\ell=0} N_{R^{mf}_{i,0}}\right\} + \sum_i \left\{\pi^h_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} \\
\mathcal{Z}_{jf} &= \sum_s \frac{\left(w_{jf}/\alpha_{sf}\right)^{-\sigma}}{\sum_{g\in m,f}\left(w_{jg}/\alpha_{sg}\right)^{-\sigma}} L_{F_{js}} - \sum_i \left\{\pi^f_{j|i} N_{R^f_i}\right\} + \sum_i \left\{\pi^w_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\}
\end{aligned}$$

Part I: Existence. $\mathcal{Z}_z = [\mathcal{Z}_{jm}, \mathcal{Z}_{jf}]$ can be decomposed into $\mathcal{Z}_z(w) = \sum_s \mathcal{L}_{sj}(w) - \sum_k \mathcal{N}_{kj}(w)$ where $\mathcal{L}_{sj}(w) \geq 0$ and $\mathcal{N}_{kj}(w) \geq 0$ are, respectively, homogeneous of degree ζ_s and $\varkappa_k$ with $\zeta_s = \varkappa_k = 0$. Since I am stacking $\mathcal{Z}_{jm}$ and $\mathcal{Z}_{jf}$, then the index z corresponds to each location. So, there is a mapping between the index z in vector $\mathcal{Z}_z$, and the index of a location j in $\mathcal{Z}_{jm}$ and $\mathcal{Z}_{jf}$:

$$j = \begin{cases} z & z \leq J \\ z - J & z \geq J+1 \end{cases}$$

First, notice that

$$\mathcal{L}_{sz}(w_m, w_f) = \begin{cases} \frac{\left(w_{jm}/\alpha_{sm}\right)^{-\sigma}}{\sum_{g\in m,f}\left(w_{jg}/\alpha_{sg}\right)^{-\sigma}} L_{F_{js}} & z \leq J \\ \frac{\left(w_{jf}/\alpha_{sf}\right)^{-\sigma}}{\sum_{g\in m,f}\left(w_{jg}/\alpha_{sg}\right)^{-\sigma}} L_{F_{js}} & z \geq J+1 \end{cases}$$

is homogeneous of degree zero. Second,

$$\mathcal{N}_{kz}(w_m, w_f) = \begin{cases} \sum_i \left\{\pi^m_{j|i} N_{R^m_i}\right\} & z \leq J, k = s \\ \sum_i \left\{\pi^h_{j|i,\ell=0} N_{R^{mf}_{i,0}}\right\} & z \leq J, k = mf0 \\ \sum_i \left\{\pi^h_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} & z \leq J, k = mf1 \\ \sum_i \left\{\pi^f_{j|i} N_{R^f_i}\right\} & z \geq J+1, k = s \\ \sum_i \left\{\pi^w_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} & z \geq J+1, k = mf1 \end{cases}$$

When $k = s$, it is easy to see that either $\sum_i \left\{\pi^m_{j|i} N_{R^m_i}\right\}$ and $\sum_i \left\{\pi^f_{j|i} N_{R^f_i}\right\}$ are homogeneous of degree zero. Similarly, when $k = mf0, \sum_i \left\{\pi^h_{j|i,\ell=0} N_{R^{mf}_{i,0}}\right\}$ is homogeneous of degree zero. Finally, by inspection,

$$\sum_i \left\{\pi^w_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} = \sum_i \left[\frac{N_{R^{mf}_{i,1}} \left(d^w_{ij}\right)^{-\theta^w}}{\sum_{l'} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{l'f}}\right)^{-\theta^{mf,m}}\right)^{\theta^w/\theta^h} \left(d^w_{il'}\right)^{-\theta^w}} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{jf}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h} \right]$$

and

$$\sum_i \left\{\pi^h_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} =$$

$$\sum_i \left[\frac{N_{R^{mf}_{i,1}} \left(d^h_{ij}\right)^{-\theta^h}}{\sum_{l'} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{l'f}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h} \left(d^{mf,f}_{il'}\right)^{-w}} \sum_l \frac{\left(d^w_{il}\right)^{-\theta^w} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{lf}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h-1}}{k^{-\theta^h}\left(w_{jm}+w_{lf}\right)^{-\theta^h}} \right]$$

are homogeneous of degree zero. Thus, we can normalize the vector w. Since Walras' Law is satisfied, and the excess demand function is continuous, existence is assured by Brouwer's fixed point theorem. $\square$

Next I discuss why the uniqueness proof fails.

Part II: Uniqueness. If $\mathcal{Z}_z$ satisfies gross substitution, that is $\frac{\partial \mathcal{Z}_z}{\partial w_z} < 0$ and $\frac{\partial \mathcal{Z}_z}{\partial w_{-z}} > 0$, where $w_z = \begin{cases} w_{jm} & z \leq J \\ w_{jf} & z \geq J+1 \end{cases}$, then uniqueness is assured. We need to compute the derivative of labor supply and labor demand to wages. Starting with labor demand, we get that:

$$\begin{aligned} \frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{1m}} &= \sum_s \frac{\partial}{\partial \log w_{1m}} \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g\in m,f} (w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1s}} \\ &= -\sigma \sum_s \left(1 - \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g\in m,f} (w_{1g}/\alpha_{sg})^{-\sigma}}\right) L_{F_{1sm}} < 0 \end{aligned}$$

And

$$\begin{aligned} \frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{1f}} &= \sum_s \frac{\partial}{\partial \log w_1} \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g\in m,f} (w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1s}} \\ &= \sigma \sum_s \frac{(w_{1f}/\alpha_{sf})^{-\sigma}}{\sum_{g\in m,f} (w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1sf}} > 0 \end{aligned}$$

And,

$$\frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{jm}} = \frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{jf}} = 0$$

Now, focusing on labor supply of single households, we get that:

$$
\begin{aligned}
\frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{1f}} &= \sum_i \frac{\pi^f_{1|i} N_{R^f_i}}{\sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}} \left(1-\pi^f_{1|i}\right)\theta^f > 0 \\
\frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{jf}} &= -\sum_i \frac{\pi^f_{1|i} N_{R^f_i}}{\sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}} \pi^f_{j|i}\theta^f < 0 \\
\frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{1m}} &= \frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{jm}} = 0
\end{aligned}
$$

For males in dual-earner households we get a similar expression. For women in dual-earner households we get that:

$$
\frac{\partial \log \sum_i \left\{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}}{\partial \log w_{1f}} = \sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}} \left(1-\pi^w_{1|i,\ell=1}\right) \sum_l \pi^h_{l|i,\ell=1,1} \frac{w_{1f}}{(w_{lm}+w_{1f})}\theta^w \geq 0
$$

$$
\frac{\partial \log \sum_i \left\{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}}{\partial \log w_{1m}} = \sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}} \left\{\pi^h_{1|i,\ell=1,1} \frac{w_{1m}}{(w_{1m}+w_{1f})} - \pi^h_{1|i,\ell=1}\tilde{\omega}_{1m}\right\}\theta^w
$$

with

$$
\tilde{\omega}_{1m} = \sum_l \frac{\pi^w_{l|i,\ell=1}\pi^h_{1|i,\ell=1,l}}{\pi^h_{1|i,\ell=1}} \frac{w_{1m}}{(w_{1m}+w_{lf})}
$$

$$
\frac{\partial \log \sum_i \left\{\pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}}{\partial \log w_{jf}} = -\sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}} \pi^w_{j|i,\ell=1} \sum_l \pi^h_{l|i,\ell=1,j} \frac{w_{jf}}{(w_{lm}+w_{jf})}\theta^w < 0
$$

$$
\frac{\partial \log \sum_i \left\{\pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}}{\partial \log w_{jm}} = \sum_i \frac{\pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{\pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}} \left\{\pi^{mf,m}_{j|i,\ell=1,1} \frac{w_{jm}}{(w_{jm}+w_{1f})} - \pi^{mf,m}_{j|i,\ell=1}\tilde{\omega}_{jm}\right\}\theta^{mf,f}
$$

with

$$
\tilde{\omega}_{jm} = \sum_l \frac{\pi^{mf,f}_{l|i,\ell=1} \cdot \pi^{mf,m}_{j|i,\ell=1,l}}{\pi^{mf,m}_{j|i,\ell=1}} \frac{w_{jm}}{(w_{jm}+w_{lf})}
$$

I stop here because it should be already clear where the gross-substitution property fails to hold. In particular, it can be shown that $\frac{\partial \log \mathcal{Z}_z}{\partial \log w_{-z}} > 0$ may not always hold. For instance, taking the derivative of excess demand for women in location 1 with respect to wages of men in location 1, that is $\frac{\partial \log \mathcal{Z}_z}{\partial \log w_{-z}}$, with $z = J+1, -z = 1$ we get:

$$\begin{aligned}
\frac{\partial \log \mathcal{Z}_z}{\partial \log w_{-z}} &= \underbrace{\sigma \sum_s \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g \in m,f} (w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1sm}}}_{>0} - \underbrace{\frac{\text{single women}_j}{\text{all women}_j} \left(\frac{\partial \log \sum_i \left\{ \pi^f_{1|i} N_{R^f_i} \right\}}{\partial \log w_{1m}} \right)}_{=0} \\
&- \underbrace{\frac{\text{married women}_j}{\text{all women}_j} \left(\sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{ \pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}} \left\{ \pi^h_{1|i,\ell=1,1} \frac{w_{1m}}{(w_{1m} + w_{1f})} - \pi^h_{1|i,\ell=1} \tilde{\omega}_{1m} \right\} \theta^w \right)}_{?}
\end{aligned}$$

The supply of women to location 1 can increase even if male wages increase when

$$\frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{ \pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}} \left(\pi^h_{1|i,\ell=1,1} \frac{w_{1m}}{(w_{1m} + w_{1f})} - \pi^h_{1|i,\ell=1} \tilde{\omega}_{1m} \right)$$

is negative enough such that the positive effect coming from other origins and from the downward-sloping demand get both out-weighted.

A3.2 Proposition 2

Proposition: *Given data on residence by gender and household type, $N_{R^k_{i,\ell}}$, employment by gender $L_{F_{jg}}$, rents r_{F_j},and commute costs d^k_{ij}, and parameters θ^k, α_s, σ_D there exists a vector of sales X_{js} and productivities A_{js} that rationalizes the observed data as an equilibrium of the model. Additionally, we get the same result if instead of employment by gender we observe employment by industry $L_{F_{js}}$ and parameters σ, α_{sg}.*

Proof. From proposition 1, we get wages. Afterwards, we obtain the total wage bill:

$$W_{js} N_{js} = w_{jm} \left[L_{F_{jm}} \right] + w_{jf} \left[L_{F_{jf}} \right]$$

Or if one observes employment by industry and parameters α_{sg} and σ:

$$W_{js} N_{js} = w_{jm} \left[\frac{(w_{jm}/\alpha_{sm})^{-\sigma}}{\sum_h (w_{jh}/\alpha_{sh})^{-\sigma}} L_{F_{js}} \right] + w_{jf} \left[\frac{(w_{jf}/\alpha_{sf})^{-\sigma}}{\sum_h (w_{jh}/\alpha_{sh})^{-\sigma}} L_{F_{js}} \right]$$

where first to second line comes from FOC: $\frac{\tilde{L}_{F_{jhs}}}{\tilde{L}_{F_{jgs}}} = \left(\frac{\alpha_{sh}/w_{jh}}{\alpha_{sg}/w_{jg}} \right)^{\sigma}$ and by summing over h and rearranging, I get $\frac{\tilde{L}_{F_{jgs}}}{\tilde{L}_{F_{js}}} = \frac{(w_{jg}/\alpha_{sg})^{-\sigma}}{\sum_h (w_{jh}/\alpha_{sh})^{-\sigma}}$.

Then, we can gets sales X_{js} from the firm's FOCs because: $\alpha_s X_{js} = W_{js} N_{js}$.

Once we obtain sales, we can recover productivity:

$$X_{js} \propto \left(\frac{W_{js}^{\alpha_s} r_{F_j}^{1-\alpha_s}}{A_{js}}\right)^{1-\sigma_D} \rightarrow A_{js} \propto \left(\frac{W_{js}^{\alpha_s} r_{F_j}^{1-\alpha_s}}{X_{js}^{\frac{1}{1-\sigma_D}}}\right)$$

□

A3.3 Proposition 3

Proposition: *Given data on residence by gender and household type, $N_{R_{i,\ell}^k}$, employment by gender $L_{F_{jg}}$, available floorspace H_i, floorspace allocation ϑ_i, rents r_{R_i} and r_{F_j} and commute costs d_{ij}^k, and parameters θ^k, η^k, ν, σ^D, α_s, β^k, there exists a vector of amenities u_i^k, productivities A_{js}, household productivity $\xi_{i\ell=0}$, sales X_{js}, floorspace wedge τ_i, and total rents E that rationalizes the observed data as an equilibrium of the model. Additionally, we get the same result if instead of employment by gender we observe employment by industry $L_{F_{js}}$ and parameters σ, α_{sg}.*

Proof. From proposition 1, we get wages. Then, we can compute residential market access:

- For single households: $\Phi_{R_i^g} = \sum_{l'} \left(\frac{d_{il'}^g}{w_{l'g}}\right)^{-\theta^g}, \forall g \in \{m, f\}$
- For couples in male breadwinner households: $\Phi_{R_{i,\ell=0}}^h = \sum_{l'} \left(\frac{d_{il}^h}{w_{lm}}\right)^{-\theta^h}$
- In dual-earner households: $\Phi_{R_{i,\ell=1,j'}}^h = \sum_l \left(\frac{d_{il}^h}{w_{lm}+w_{j'f}}\right)^{-\theta^h}$ $\Phi_{R_{i,\ell=1}}^w = \sum_l \left(\frac{d_{il}^w}{\left(\Phi_{R_{i,\ell=1,l}}^h\right)^{1/\theta^h}}\right)^{-\theta^w}$
- On average:

$$\Phi_{R_i^{mf}} = \left\{\left[\frac{1}{\gamma^w}\left(\Phi_{R_{i,\ell=1}^w}\right)^{-1/\theta^w}\right]^{-\nu} + \left[(\xi_{i\ell=0})^{-1}\left(\Phi_{R_{i,\ell=0}^h}\right)^{-1/\theta^h}\right]^{-\nu}\right\}$$

where $\gamma^w = \Gamma\left(\frac{\theta^w-1}{\theta^w}\right)(T^w)^{1/\theta^w} = 1$

Using data on labor force participation choices, we can use the following equation to recover household productivities:

$$\frac{\mu_i^{mf}}{1-\mu_i^{mf}} = \left(\frac{\left(\Phi_{R_{i,\ell=1}}^w\right)^{-1/\theta^w}}{(\xi_{i\ell=0})^{-1}\left(\Phi_{R_{i,\ell=0}^h}\right)^{-1/\theta^h}}\right)^{-\nu}$$

i,e,

$$\xi_{i\ell=0} = \frac{\left(\Phi_{R_{i,\ell=1}}^w\right)^{1/\theta^w}}{\left(\Phi_{R_{i,\ell=0}^h}\right)^{1/\theta^h}}\left(\frac{\mu_i^{mf}}{1-\mu_i^{mf}}\right)^{\frac{1}{-\nu}}$$

Then using the residential supply conditions, from which we can recover implied amenities once that we know wages and rents:

$$
\begin{aligned}
N_i^g \equiv \pi_i^g N^g &= \sum_l \left(\frac{r_{R_l}^{1-\beta^g}}{u_l^g \Phi_{R_l^g}^{1/\theta^g}} \right)^{-\eta^g} N^g \left(u_i^g \Phi_{R_i^g}^{1/\theta^g} r_{R_i}^{\beta-1} \right)^{\eta^g} \\
N_i^{mf} \equiv \pi_i^{mf} N^{mf} &= \sum_l \left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_l^{mf} \Phi_{R_l^{mf}}^{1/\nu}} \right)^{-\eta^{mf}} N^{mf} \left(u_i^{mf} \Phi_{R_i^{mf}}^{1/\nu} r_{R_i}^{\beta-1} \right)^{\eta^{mf}}
\end{aligned}
$$

Then we need to solve for unobservables on the housing side. First, we need to get back sales using proposition 2. Then, it is needed to introduce a new pair of location characteristics since the floorspace market clearing condition $E_{Ri} = r_{R_i} H_{R_i}$ will not necessarily hold at the values for data and wages. Therefore, we need to introduce an additional unobservable that can be interpreted as quality of housing $\tilde{H}_{R_i} = H_{R_i} q_{R_i}$ where H_{R_i} are the physical units of floorspace. From the housing market clearing condition we get that $E_{Ri} = r_{R_i} \tilde{H}_{R_i} = r_{R_i} H_{R_i} q_{R_i} \to q_{R_i} = \frac{E_{Ri}}{r_{R_i} H_{R_i}}$. Similar residuals can be defined for commercial floorspace: $q_{F_i} = \frac{X_i}{r_{F_i} H_{Fi}}$.

Finally, we just need to solve for land use wedge, which can be identified from:

$$
(1 - \tau_i) = \frac{r_{R_i}}{r_{F_i}}
$$

for locations with mixed land use. For locations with single land use, these wedges cannot be identified but are rationalized by zero productivities for all sectors or zero amenities for all worker groups. $\square$

B Appendix Tables and Figures

B.1 Tables

Table B.1: Commuting and Wages

	Husband in D.E. Households	Wife in D.E. Households
	(1)	(2)
Wage at one's destination	2.5647	1.6135
	(0.3202)***	(0.1699)***
Wage at one's destination ×	-0.4466	-0.2744
HH income at destination pair	(0.0588)***	(0.0297)***
Origin-spouses' destination FE	X	X
Destination-pair FE		
Min Elasticity	1.856	1.178
St.Dev	(0.2288)	(0.1242)
Max Elasticity	3.054	1.914
St.Dev	(0.3839)	(0.2019)
N	90477	94773

Table B.2: Differences-in-Differences of New Transit Infrastructure on Mobility

Panel A:	Total Population	Singles to Married	Females to Males
	(1)	(2)	(3)
$Treat \times (Year = 2017)$	-0.0964	0.0477	-0.0041
	(0.0212)***	(0.0137)***	(0.0030)
Panel B:	Single to Married (Males)	Single to Married (Females)	Dual-Earners to Male Breadwinner
	(1)	(2)	(3)
$Treat \times (Year = 2017)$	0.0534	0.0418	-0.0070
	(0.0147)***	(0.0142)***	(0.0135)
Block FE	X	X	X
Year FE	X	X	X
N	1190	1190	1190

Table B.3: The Effect of the New Transit Infrastructure - Mechanisms

	Full	No Externalities	No Mobility	No LFP	Neither of the three
	(1)	(2)	(3)	(4)	(5)
Real Income	0.0107	0.0139	0.0062	0.0096	0.0075
Real Income per Worker	0.0070	0.0102	0.0015	0.0099	0.0075
Gender Gap	-0.0215	-0.0200	-0.0233	-0.0112	-0.0114
Gender Gap (per Worker)	-0.0105	-0.0084	-0.0078	-0.0159	-0.0149
Gender Gap, Singles	0.0048	0.0089	0.0083	-0.0018	0.0058
Gender Gap (per Worker), Singles	0.0056	0.0104	0.0098	-0.0022	0.0068
Gender Gap, Married	-0.0413	-0.0401	-0.0457	-0.0200	-0.0227
Gender Gap (per Worker), Married	-0.0254	-0.0237	-0.0230	-0.0285	-0.0303
# DE/# Couples	0.0174	0.0174	0.0223	0	0

B.2 Figures

Figure B.1: The Line 1 of Lima's Metro

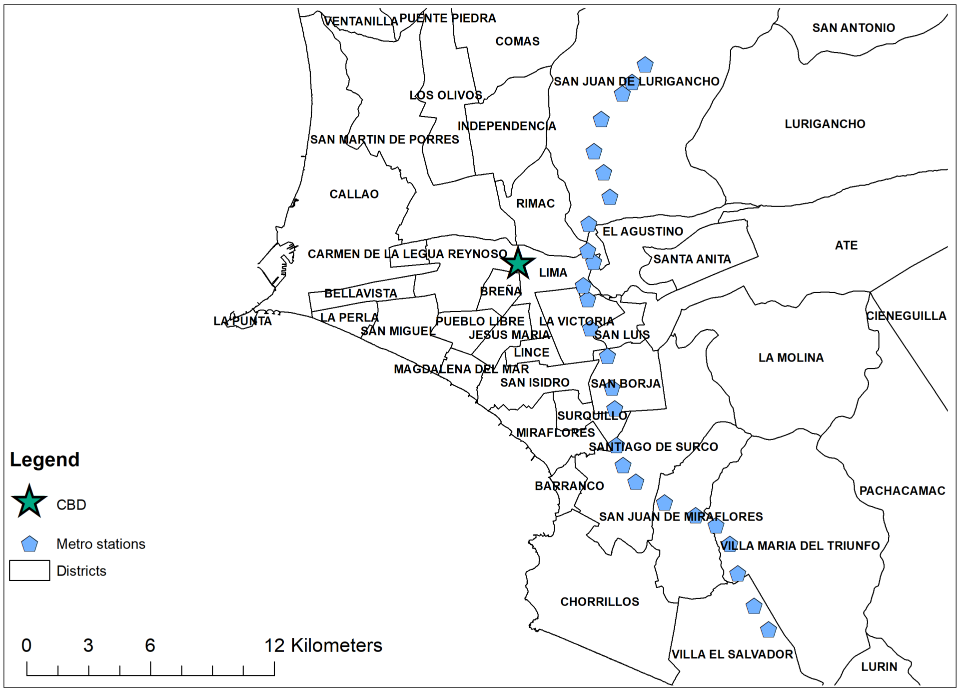


Figure B.2: The Bus Rapid Transit System

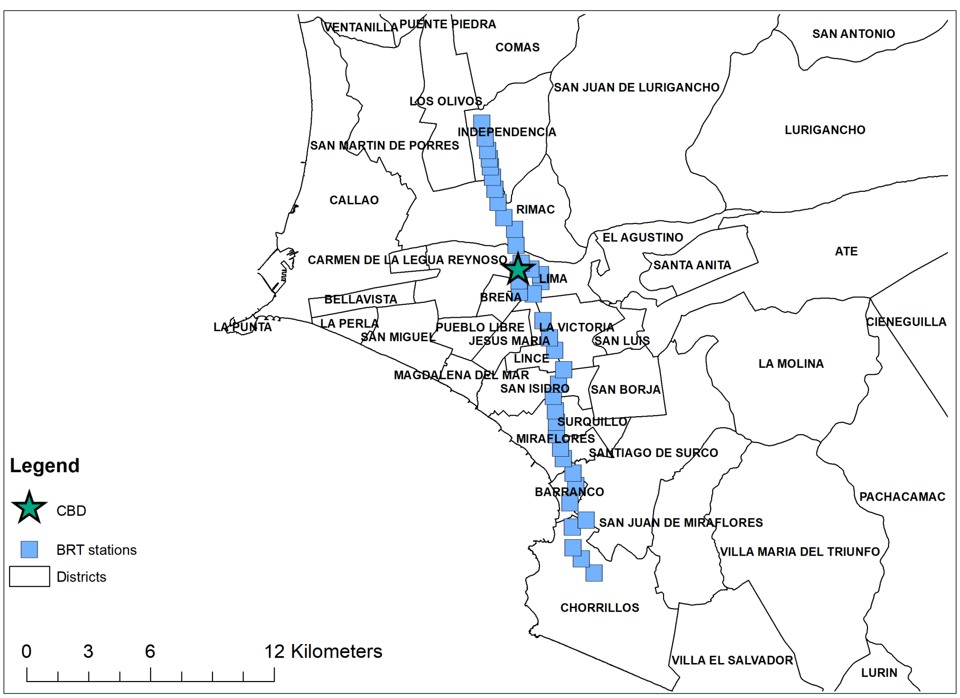
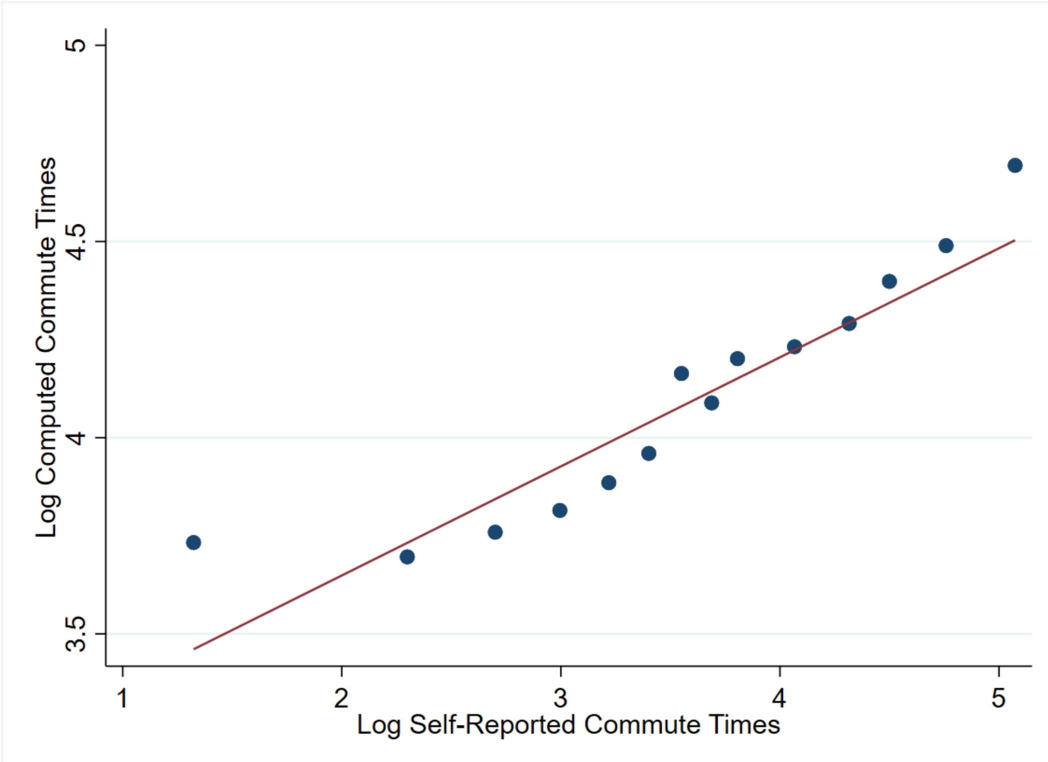


Figure B.3: Imputed versus reported commuting times



Self-reported commute times, at the district level, come from the 2010-2018 waves of the Commuting Surveys. Imputed commute times come from the Open Street Road Network data in par with Imputed speeds. Imputed commute times were computed at the zone level, and then aggregated to the district level using median values. Scatter was generated using the *binscatter* command in Stata. The correlation between the two measures of commute times is of 53%.

Figure B.4: The Reduced Form Impact of Transit Infrastructure on Rents, Income, and Extreme Poverty

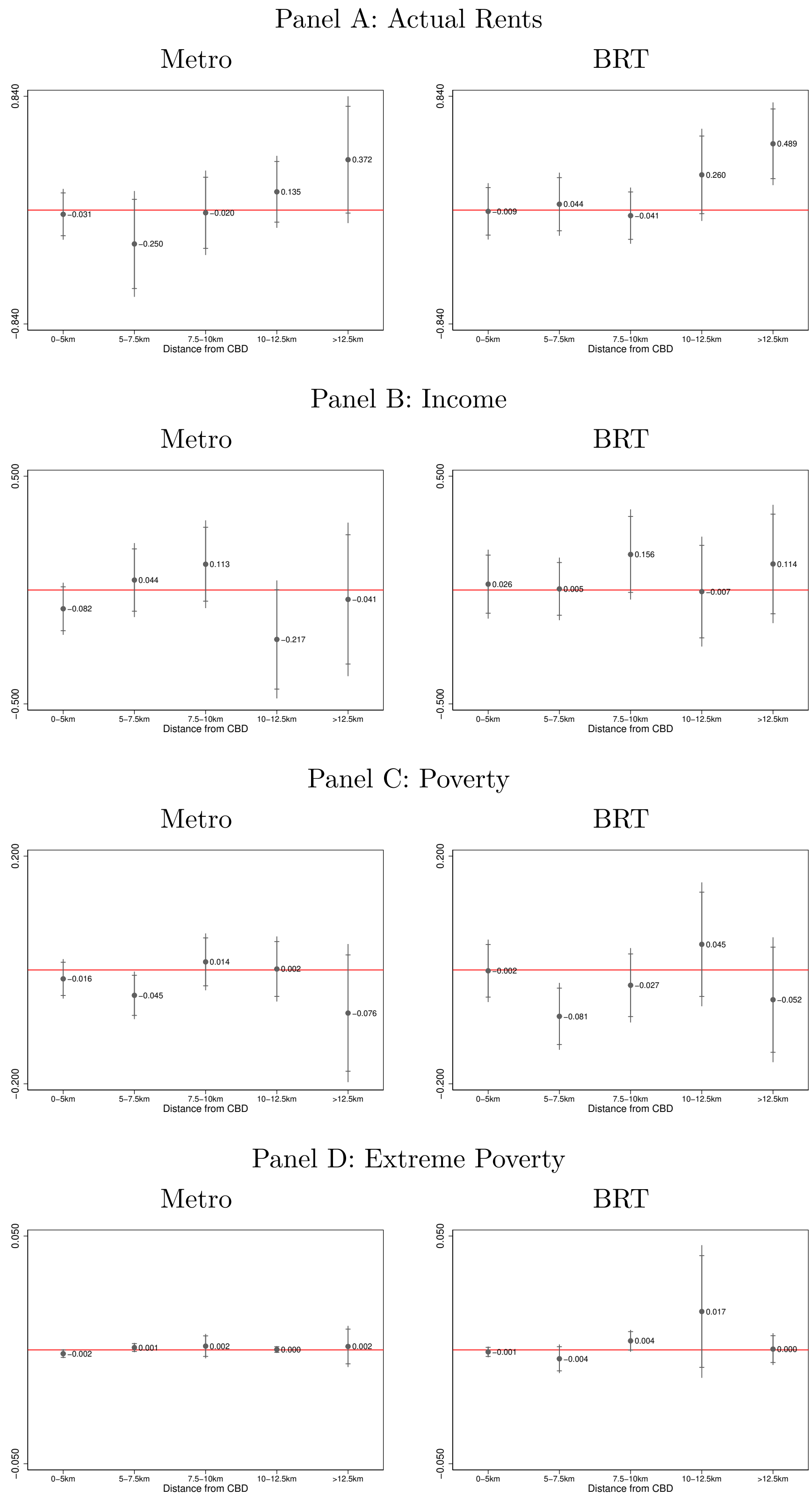


Figure B.5: Timing Assumptions

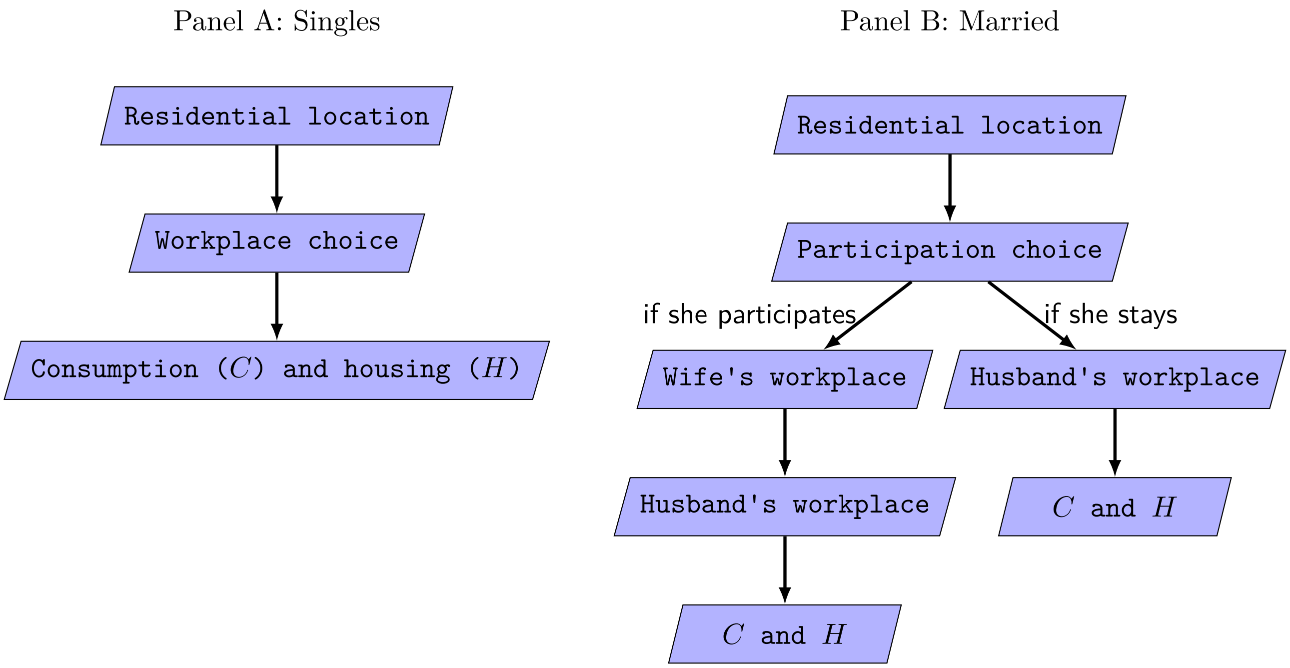
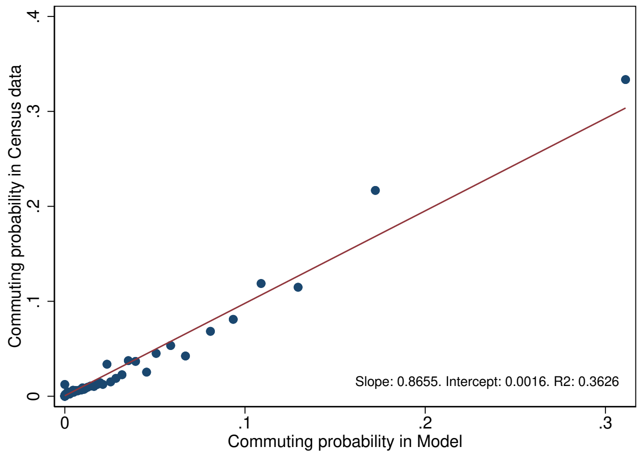


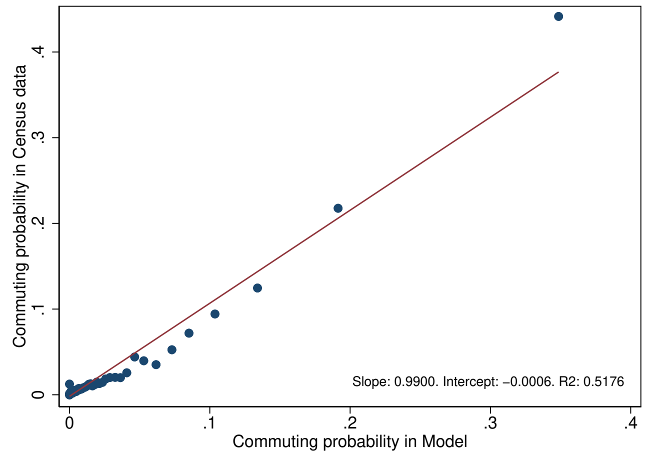
Figure B.6: Commuting and Over-identification

Panel A: Females in Dual-Earner Households

Independent Commute

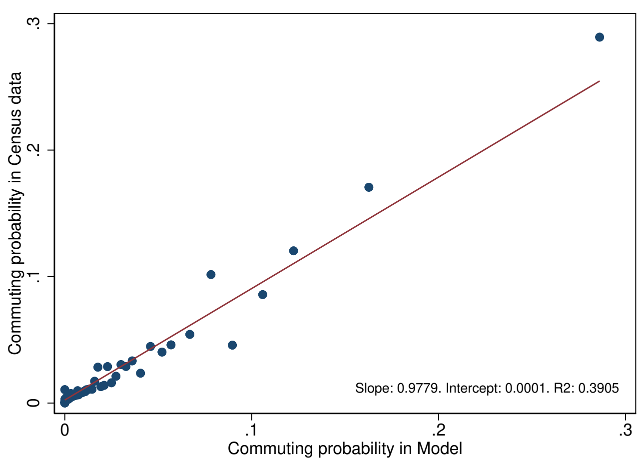


Interdependent Commute



Panel B: Males in Dual-Earner Households

Independent Commmte



Interdependent Commute

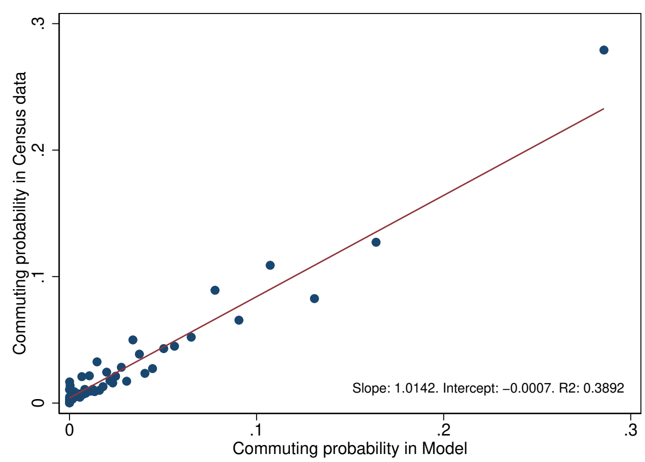


Figure B.7: The Reduced-form Impact of Transit Infrastructure, Data and Model

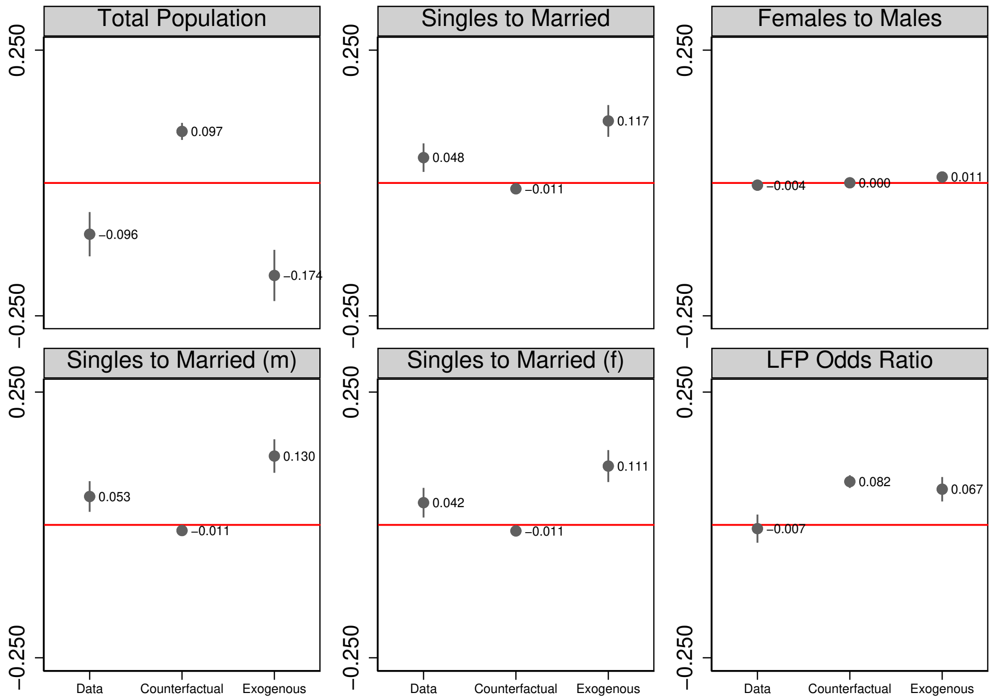


Figure B.8: Reduced-form Impact by Distance to the CBD, Data and Model

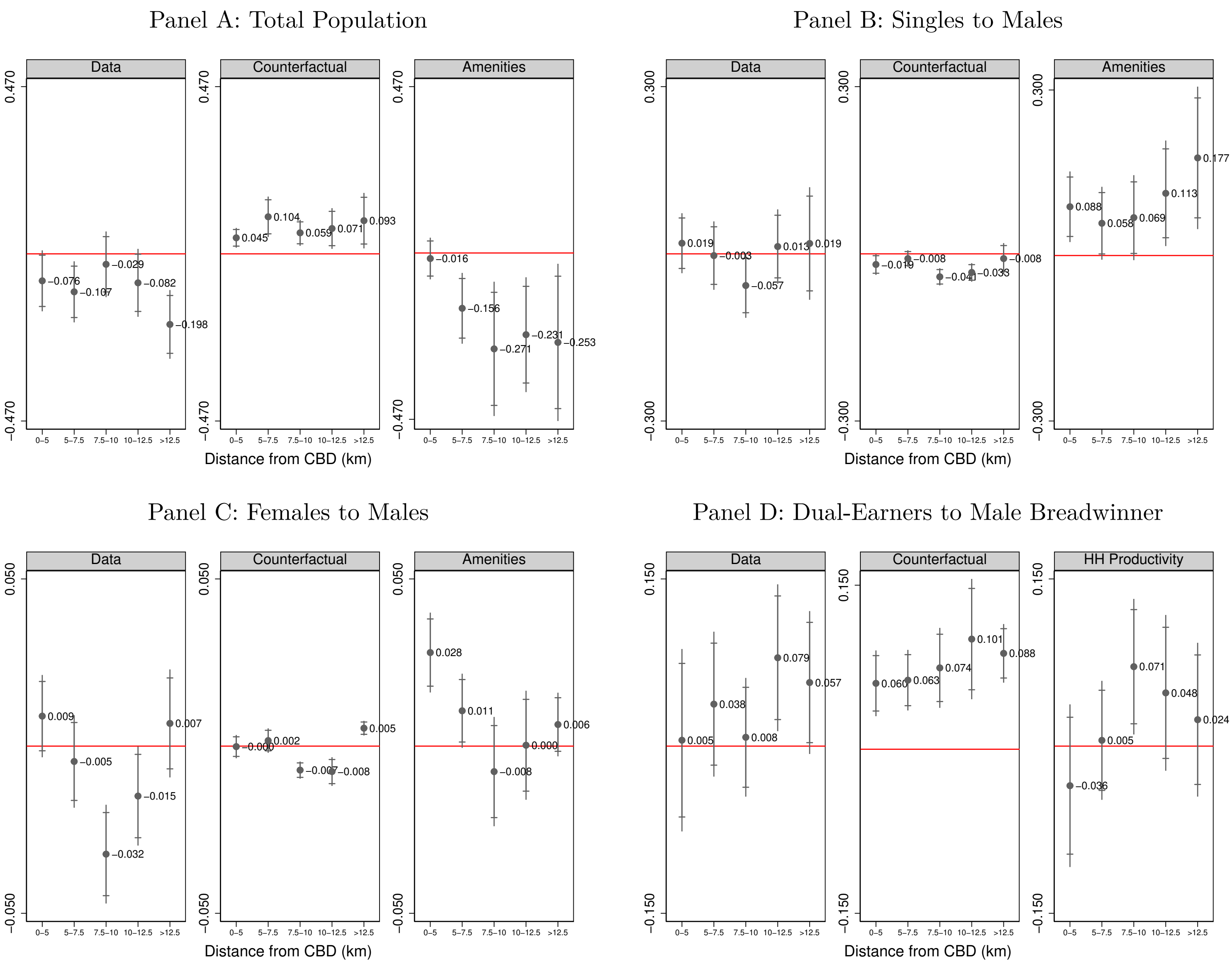
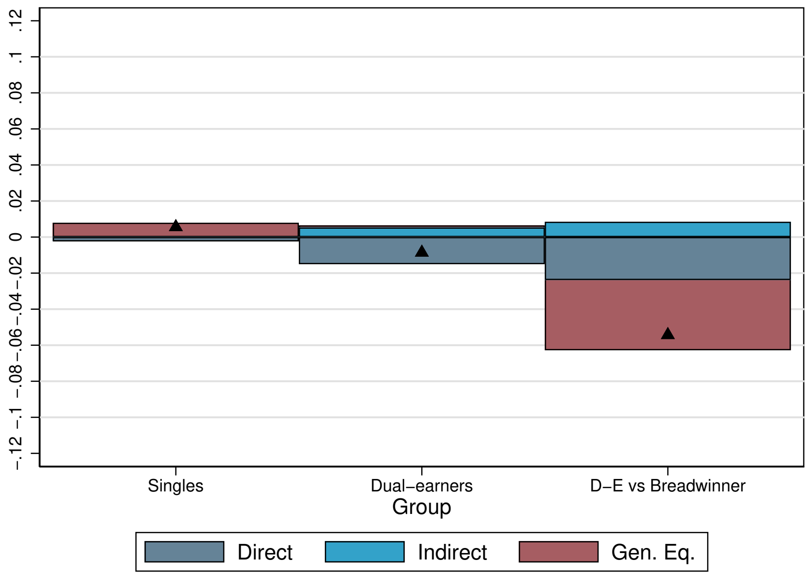


Figure B.9: Decomposition of the Aggregate Impacts

Panel A: Incremental Impact of Metro Lines 2 and 4



Panel B: Overall Impact of Metro Lines (1, 2, 4) and the BRT