Banco de México

Documentos de Investigación

Banco de México

Working Papers

N° 2016-19

Insurance Against Local Productivity Shocks: Evidence from Commuters in Mexico

Fernando Pérez-Cervantes
Banco de México

November 2016

Documento de Investigación
2016-19

Working Paper
2016-19

Insurance Against Local Productivity Shocks: Evidence from Commuters in Mexico*

Fernando Pérez-Cervantes[†]
Banco de México

Abstract: I slightly modify the model of Monte et al. (2015) to estimate how workers in Mexican municipalities choose the location of their workplace based on the income gains from commuting to another municipality. Estimates are in line with the intuition: Static estimates for both 2010 and 2015 suggest that those who commute earn an average 30 percent more than their non-commuting counterparts, and that commutes tend to be to municipalities located close to the place of residence. Comparing both years suggests that a reduction in local productivity both decreases the number of workers that come from other municipalities and increases the number of local residents that decide to work somewhere else, mitigating the negative effect of the reduction in local wages with higher earnings from the new work destinations. I find that some municipalities were not able to mitigate the negative productivity shocks on their income.
Keywords: Commuting, Economic Geography, Mexico
JEL Classification: F1, R1, J6, O2

Resumen: Modifico ligeramente el modelo de Monte y coautores (2015) para estimar cómo los trabajadores en municipios mexicanos escogen el lugar de trabajo con base en los aumentos a su ingreso provenientes de trasladarse a otro municipio. Las estimaciones están acorde con la intuición: Estimaciones estáticas para 2010 y 2015 sugieren que los trabajadores que viajan a otros municipios ganan en promedio 30 por ciento más que sus contrapartes que no lo hacen, y que los traslados tienden a ser a los municipios que se ubican cerca del lugar de residencia. La comparación entre estos dos años sugiere que una reducción en la productividad local disminuye el número de trabajadores que llegan de otros municipios y aumenta la cantidad de los residentes locales que deciden trabajar en otro lugar, así mitigando el efecto negativo de la reducción en los salarios locales a través de los ingresos más altos provenientes de nuevos lugares de trabajo. Encuentro que algunos municipios no lograron mitigar los choques negativos de productividad sobre su ingreso.
Palabras Clave: Traslado de trabajadores, geografía económica, México

---

*Special thanks to Alfonso Cebreros, who came up with the idea of the whole thing being an insurance mechanism and then a lightning struck me. Thanks to Óscar Cuéllar for excellent research assistance.
[†] Dirección General de Investigación Económica. Email: fernando.perez@banxico.org.mx.

1 Introduction

In 2015, around 8 million salaried workers in Mexico had a job in a different municipality from the one where they had their residence, or 19.6 percent of the total salaried workers of the country. These workers, which will be named intermunicipal commuters (IMC) henceforth, commuted a self-reported median of between 31 and 60 minutes every working day, each way (between 5 and 10 hours per week), according to the *Encuesta Intercensal 2015*.[1] In 2010, the recently published *Muestra Especial del Censo de Población y Vivienda 2010* shows that there were 7.3 million IMC, or 19.2 percent of the salaried workers.[2]

Both data sets show that IMC earn on average around 30 percent more than the average salary of their municipality of residence, and that workers who arrive to a certain municipality earn just over 5 percent more than the average salary of their destination.[3] This suggests that salaries seem more inherent to the jobs at the municipality than to the workers themselves meaning that, if they were taking the location of residence as fixed, workers look for higher-than-locally-sourced wages across distant but feasible locations, with the trade-off of having to take a long commute to earn this higher wage.

Using this last fact, in this document I work on a model where workers take location of residence as fixed, and decide where to work. This is a completely opposite view of the traditional *town center* model, where location of work is fixed and it is the location of

[1] The place of residence is defined as the location where the person sleeps at least 182 nights of the year, and the place of work is defined as the location where the person worked the week before the Census.

[2] There is no self-reported data on commuting time for 2010.

[3] This results holds for both 2010 and 2015.

residence the one that is chosen.[4] This document makes use of a slightly modified version of the model in Monte, Redding, and Rossi-Hansberg (2015),[5] the two data sets mentioned in the first paragraph, and the *Red Nacional de Caminos* to calculate the changes in income and productivity at the municipal level and separates the change in municipal income into the contribution of the changes in their own local productivity and the contribution of the changes in the productivity of the municipalities where the residents decide to work in. This way I identify which municipalities were able to distribute the productivity shocks better (by sending less workers to the municipalities that had negative shocks, and more workers to the municipalities with positive shocks). Finally, I identify which municipalities could not adjust over this margin.

This large number of commuters is informative of why the usual surveys of income that policymakers in Mexico use to tackle important issues such as poverty and employment (examples are ENIGH, ENEU, etc.) are not used to study things like local wages and productivity: the wage is labeled to the person and place of residence, not to the place of work. It is impossible to draw plausible conclusions about productivity assuming the income is sourced locally.[6] This paper tangentially relates to this issue, and looks at the positive side of the problem: what can we learn from the commuters' decisions, assuming

[4]Alas, town center models are used to study where and why do people choose the location of their residence.

[5]In their model, workers choose the location of work and of residence simultaneously. This document studies the decision of commuting to increase wages. Neither the original model or my simplification of taking the location of residence as given allow to pin down the 5% average wage differential between non commuters and incoming workers to a municipality. Models that study location of residence and variations in workplace wage must take into consideration a general concept of amenities and self-selection, which is not studied in this document.

[6]See Pérez-Cervantes (2013) for an extensive discussion on how can productivity be incorrectly measured with residents' income instead of workers' income.

they are optimal? These foundations lead the way into how to correctly answer questions about local productivity, as well as measure local workers' exposure to trade with other municipalities, resident exposure to income from other municipalities, and so on. There were hundreds of municipalities in Mexico whose income from wages came primarily from other municipalities in 2010 and 2015.[7] Most likely this is still the case. Public policy aiming for increases in the income of the poorest municipalities may not necessarily mean that they should be local policies: they might include improving transportation infrastructure to the municipalities where there are higher wages, or policies aiming to increase productivity in the municipalities where most of the income is sourced from in poor municipalities.

Data from both 2010 and 2015 indicate that 1 out of 5 workers get their income from another municipality so concluding that, for example, a municipality saw a reduction in poverty during those 5 years because of something that happened locally can lead to enormous mistakes in diagnosis and praxis for policy. I address this problem by separating the municipalities that saw an increase in income despite a reduction in productivity because of increases in commutes from the ones that saw both a reduction in productivity and income. It turns out that the latter municipalities were not able to produce commuters, thus could not mitigate the negative shock.

This paper is organized as follows: section 2 shows the model that is used to estimates the gains from commuting and section 3 describes the data used for the paper, with a special subsection on describing how was the data averaged out to calibrate the model. Section 4 explains the results, and section 5 concludes.

[7] See figure 3 for the case of 2015.

2 Model

I use a model of monopolistic competition, slightly modifying the one in Monte, Redding, and Rossi-Hansberg (2015). The modification is that in this model I let the place of residence fixed.[8] This model is used because it is a good theoretical framework that is able to separate the sources of municipal payments to the factors of production into three parts that are not necessarily independent, but that have historically all been associated with the term "productivity". The three terms are (a) the number of workers in a municipality,[9] (b) the geographic characteristics of the municipality,[10] and (c) marginal cost.[11] In my model, I call productivity only to the third component, so it is possible to compensate for the huge variation in the number of workers and in geographic characteristics observed in the data (there's municipalities with close to a million workers, and there are municipalities with just over a hundred), without forcing the model to give some municipality 1000 times more *average* productivity than others, which is quite unrealistic. Added to the tractability of the model, the model is convenient for policy simulations: it is possible to get the changes in the municipal outcomes by shocking the three components separately, and see that municipalities change very differently depending on which one of the components is shocked, making the better policy to be regionally dependent. No other model, to the best of my knowledge, can

[8]In the cited paper, workers jointly chose the pair of location of residence and location of workplace.

[9]Regions with high productivity attract more workers, thus reducing wages. Two regions with free labor mobility and the same wages but one with the first one having more workers than the second will most likely have a larger productivity in the first.

[10]Transport costs act as a trade barrier, so regions that face the lowest transport cost can use more productive resources in the production line rather than in transportation, therefore allowing to have more output per unit of endowment.

[11]Two regions with the same wages but the first one with lower marginal costs than the second one most likely have the first region being more productive than the second one.

do this separation for a large number of regions, which is the case of this document.

Another level of complexity that is tackled with the model is taking it to the data. I want to study wages and productivities at the municipal level in Mexico, but I also want the commuting patterns of the workers who are accepting those wages. The only data set that is publicly available and has a large number of pairs of municipal locations of residents and workers that enables me to do this is the Mexican Census data, which contains only labor data, but representative of every worker in each municipality in the country. So it is labor productivity and variations of labor productivity that will be estimated and be used as one of the explanatory variables for changes in the patters of commuting at the municipal level. This productivity, which algebraically is just a marginal cost for unit wages, can be seen as the combination of other factors of production such as capital and land. Since this is still given for the worker when making the commuting decision, and all the unobserved factors and their returns are part of the revenues of the municipality where the output is produced, the productivity parameter in this model is in fact just a relabel for the rest of the production function of a monopolistic competition model with gravity as the source of trade, which is the case of the one used in this paper.[12]

This document will have an economy that consists of $N = 2456$ municipalities (indexed i, j, or n on the rest of the paper, and each representing one of the municipalities of Mexico)[13] where the only factor of production is labor; trade in goods market happens for an iceberg

[12] In a Cobb-Douglass setting, productivity would be a combination of TFP and capital per worker, $A\left(\frac{K}{L}\right)^{\alpha}$.

[13] There were 2456 municipalities in Mexico in 2010. Then in 2011 a municipality called Othón P. Blanco in Quintana Roo split into two municipalities. In order to make comparisons with 2015, all the 2015 data for these two municipalities will be merged as if it was still one municipality, with all its economic activity located in Othón P. Blanco.

cost, and labor can commute from the municipality of residence to the municipality of work, also for a cost, which will be modeled as a non-studied residual: the the mean non-pecuniary aspects of commuting, or amenities.[14]

2.1 Workers and Residents (consumers)

Residents (consumers) in municipality n are endowed with one unit of labor and work in municipality i for a wage w_i (there is no restriction whether or not n can be equal to i) that they use in its entirety to purchase a measure M of varieties ω of consumption goods. This measure of varieties is available to any resident of any municipality (although they might have to pay a different price for it depending on the municipality where they live), and consumers combine the varieties in a CES aggregator. The utility derived for the consumption of the CES aggregator of a worker with identity ϕ in municipality i that resides in municipality n will be multiplied by two numbers in order to get the total utility of each individual worker. The first one is an idiosyncratic realization of a random variable $b_{in}(\phi) \sim Frechet\left(\left(\Gamma\left(\frac{\lambda-1}{\lambda}\right)\right)^{-\lambda}, \lambda\right)$ with $\lambda > 1$.[15] The second number is $(B_{in})^{\frac{1}{\lambda}} > 0$, the mean non-pecuniary aspects of commuting from municipality n to i.[16] So, large values of B_{in} imply that, conditional on living in n, working in i is more desirable, *despite* the wage. The

[14]See Pérez-Cervantes and Cuéllar (2016) to see how this residual can be studied in detail.

[15]If $\lambda < 1$, the expected value of this random variable is not defined. If $\lambda > 1$, the expected value of this random variable is 1. The function $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is positive, strictly increasing, and finite for $x \in [0,\infty)$.

[16]Every worker ϕ living in n will have 2456 iid realizations of $b_{in}(\phi)$, namely $\{b_{1n}(\phi), b_{2n}(\phi), ..., b_{2456}(\phi)\}$, and this assumption allows to identify the average non-pecuniary aspects of commuting, while at the same time allowing the workers to rationally decide to go to places with low wages and low amenities (albeit with low probability).

total utility of a worker with identity ϕ in municipality i that resides in municipality n will be

$$U_{in\phi} = (B_{in})^{\frac{1}{\lambda}} b_{in}(\phi) \left(\int_0^M c_n(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \tag{1}$$

Each worker of i living in n with identity ϕ takes the price of the varieties as given and has the same wage (which will not be a function of ϕ)[17] and so the budget constraint is identical for all of them, independent of the realization of the random variable $b_{in}(\phi)$

$$\int_0^M p_n(\omega) c_n(\omega) d\omega \leq w_i \tag{2}$$

The solution for this problem is standard, and the demand for every variety ω is

$$c_n(\omega) = \frac{(p_n(\omega))^{-\sigma}}{(p_n)^{1-\sigma}} w_i \tag{3}$$

where p_n is the ideal price index of consumption in municipality n, and satisfies

$$(p_n)^{1-\sigma} = \int_0^M (p_n(\omega))^{1-\sigma} d\omega \tag{4}$$

So the last series of equations and definitions allow for all workers in municipality i to have the same wage w_i regardless of where they live, and all consumption in municipality n has the same price index p_n regardless of where do residents work, so workers living in municipality i and working in municipality n they all have the same budget constraint and will make the

[17] See Pérez-Cervantes and Cuéllar (2016), where wages depend on observable characteristics of individuals such as the age, education and profession.

same consumption choices.[18]

Finally, all the consumption in the same municipality will be done in the same proportions, and will only vary by a constant: the wage.[19] In equilibrium, indirect utility of resident with identity ϕ of municipality n working in i will be just the real wage multiplied by the realization of the random variable $b_{in}(\phi)$ and by $(B_{in})^{\frac{1}{\lambda}}$

$$U_{in\phi} = \frac{(B_{in})^{\frac{1}{\lambda}} b_{in}(\phi) w_i}{p_n} \tag{5}$$

This means indirect utility for worker ϕ will vary depending on the location of work, because the municipality that is going to be chosen as location of work by each resident of municipality n will be the one that gives every individual its maximal utility

$$i = \arg\max_j \left\{ \frac{(B_{jn})^{\frac{1}{\lambda}} b_{jn}(\phi) w_j}{p_n} \right\} \tag{6}$$

Now, a little of useful algebra. If $b_{jn}(\phi) \sim Frechet\left(\left(\Gamma\left(\frac{\lambda-1}{\lambda}\right)\right)^{-\lambda}, \lambda\right)$ then $\left(B_{jn}\right)^{\frac{1}{\lambda}} b_{jn}(\phi) \sim Frechet\left(\left(\Gamma\left(\frac{\lambda-1}{\lambda}\right)\right)^{-\lambda} B_{jn}, \lambda\right)$.[20] Then, a worker ϕ living in municipality n will pick municipality i because it is the largest utility provider with the following

[18]This simplification can be thought of as consumption not requiring any time, or that workers do not have a better technology to transport final goods than the producers, so they just the consumption goods in the municipality of residence.

[19]This feature comes from the Dixit-Stiglitz preferences.

[20]Proof: Let $X \sim Frechet(T,\lambda)$. This implies $\Pr(X \leq x) = \exp\left(-Tx^{-\lambda}\right)$. Define $Y = \alpha X$ and so $\Pr(Y \leq y) = \Pr(\alpha X \leq y) = \Pr\left(X \leq \frac{y}{\alpha}\right) = \exp\left(-T\left(\frac{y}{\alpha}\right)^{-\lambda}\right) = \exp\left(-\left(T\alpha^{\lambda}\right) y^{-\lambda}\right)$ which means $Y \sim Frechet(T\alpha^{\lambda}, \lambda)$. Q.E.D.

probability:

$$\Pr\left(\frac{(B_{in})^{\frac{1}{\lambda}}\, b_{in}(\phi)\, w_i}{p_n} \geq \max_{j\neq i} \frac{(B_{jn})^{\frac{1}{\lambda}}\, b_{jn}(\phi)\, w_j}{p_n}\right) = \frac{B_{in} w_i^{\lambda}}{\sum_{j=1}^{2456} B_{jn} w_j^{\lambda}} \tag{7}$$

Note that this probability has intuitive properties, such as it being increasing in the wage of the municipality in the numerator (i.e. the municipality whose probability of being the highest utility provider is being measured), and decreasing in the others' wages. Note the important role of the non-pecuniary aspects of commuting, too. If they are large, they attract workers, even if municipalities offer a low wage. If they are small (or even equal to zero), then municipality i won't attract workers from n even if they offer large wages.[21] I then invoke the Law of Large numbers and will assume that the proportion of workers living in municipality n that will pick municipality i is exactly the probability in equation 7.

2.2 Producers

Following standard literature,[22] every variety $\omega \in [0, M]$ is produced by only one municipality under monopolistic competition. To produce a variety ω, a firm in municipality i hires labor and uses the first $F > 0$ units of labor as fixed costs, and to be able to produce this firm faces constant marginal cost $1/A_i$ per unit of labor, indicating the municipality's inverse of the productivity (every firm in the same municipality has the same productivity). Assume without loss of generality that every variety ω produced in i are in a convex set, $\omega \in [\underline{M_i}, \overline{M_i}] \subset [0, M]$,

[21] Pérez-Cervantes and Cuéllar (2016) study B_{in} as a function of the commuting time between i and n as well as some other factors.

[22] See Krugman (1991) and Helpman (1998).

that is, the varieties produced in municipality i are all identified together, irrespective of what the actual final product is.[23] The labor requirement in region i to produce $x_i(\omega)$ units of variety ω is therefore

$$\ell_i(\omega) = F + \frac{x_i(\omega)}{A_i} \tag{8}$$

The standard profit maximization problem implies that the producer price of variety ω in municipality i, defined as $p_{ii}(\omega)$ is

$$p_{ii}(\omega) = \left(\frac{\sigma}{\sigma - 1}\right)\frac{w_i}{A_i} \tag{9}$$

which implies that for every $\omega \in [\underline{M_i}, \overline{M_i}]$ the producer price is the same,[24] and the zero profit condition implies that production is also the same for every ω, and equal to

$$x_i(\omega) = F(\sigma - 1)A_i \tag{10}$$

Assigning $\delta_{ni} \geq 1$ the iceberg cost of taking good ω from municipality i to n, then, defining $p_{ii}(\omega)$ to be the producer price in municipality i we have that the price for the same variety in any other municipality is

$$p_{ni}(\omega) = \delta_{ni} p_{ii}(\omega) = \left(\frac{\sigma}{\sigma - 1}\right)\frac{\delta_{ni} w_i}{A_i} \tag{11}$$

[23]This is a simplification of notation, justified because the utility function assigns equal weights to all varieties ω, which is the case in this paper. Otherwise instead of intervals we would have just abstract sets.

[24]This result concludes the justification that the varieties ω can be ordered by municipality of production.

In equilibrium, same producer price and equal weights in the utility function imply that for every $\omega \in [\underline{M_i}, \overline{M_i}]$, the same number of workers is hired by every firm in the same municipality, namely

$$L_i^W(\omega) = F\sigma \tag{12}$$

Notice this number is not only constant for every $\omega \in [\underline{M_i}, \overline{M_i}]$ but also for every i.[25] The total number of varieties produced in municipality i is

$$\overline{M_i} - \underline{M_i} = \frac{L_i^W}{F\sigma} \tag{13}$$

where L_i^W is the total number of workers in municipality i, independent of where they live.

Ordering the varieties ω such that all the ones coming from the same municipality, and making $\omega \in [\underline{M_i}, \overline{M_i}]$ the set of varieties ω that are produced in municipality i implies the expression for consumption ends up being

$$c_n = \left(\sum_{i=1}^{2456} \int_{\underline{M_i}}^{\overline{M_i}} c_{ni}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \tag{14}$$

with $\underline{M_1} = 0$, $\overline{M_i} = \underline{M_{i+1}}$, and $\overline{M_{2456}} = M$. Also, p_n is the price index in municipality n

$$p_n = \left(\sum_{i=1}^{2456} \int_{\underline{M_i}}^{\overline{M_i}} p_{ni}(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}} \tag{15}$$

[25]This assumes a constant fixed cost F for every good in every municipality. This assumption will not be relaxed in this model, but F can be thought of as the proportion of workers in every firm that are not in the production line but nevertheless are necessary for production to occur, such as managers or lawyers.

2.3 Trade

With equal weights in the utility function and constant prices of goods coming from every municipality, we have a very simple expression for the fraction of municipality n's expenditure in goods coming from i. This fraction is

$$\pi_{ni} = \frac{\left(\overline{M_i} - \underline{M_i}\right) p_{ni}^{1-\sigma}}{\sum_{j=1}^{2456} \left(\overline{M_j} - \underline{M_j}\right) p_{nj}^{1-\sigma}} \tag{16}$$

which is a very common gravity equation. Substituting the values for measure of product varieties as a function of the total number of workers and substituting for producer prices, we get rid of the fixed cost F, of the markups and of the number of varieties, and obtain an expression that has information that can be obtained from the data, namely labor demand, productivity, transport cost, and wages. This expression for the trade share becomes

$$\pi_{ni} = \frac{L_i^W A_i^{\sigma-1} \left(\delta_{ni} w_i\right)^{1-\sigma}}{\sum_{j=1}^{2456} L_j^W A_j^{\sigma-1} \left(\delta_{nj} w_j\right)^{1-\sigma}} \tag{17}$$

In fact, the term $L_i^W A_i^{\sigma-1}$ which usually is the productivity in Eaton and Kortum (2002) types of models (EK henceforth), now governs the proportion of sales with two forces: one equivalent to the EK productivity (with a relative importance of $(\sigma-1)/\sigma$) and another coming from the labor supply (with a relative importance of $1/\sigma$).

2.4 Equilibrium

In equilibrium, the endowment L_n^R of residents in every municipality n is distributed to the rest of the municipalities i to work according to some probability distribution $\rho(i|n)$,[26] giving a resulting number of workers in every municipality $\sum_{n=1}^{2456}\rho(i|n)L_n^R$ equal to the demand L_i^W. Notice that this same probability distribution gives the total income Y_n^R of municipality n, equal to the weighted sum of the income coming from all the municipalities

$$Y_n^R = \sum_{i=1}^{2456}\rho(i|n)\,w_i L_n^R = L_n^R\sum_{i=1}^{2456}\rho(i|n)\,w_i = L_n^R v_n \tag{18}$$

where v_n is the average resident wage in municipality n, no matter where they work. The number of workers in every municipality, the iceberg costs, and the productivities all interact in equation 17, so imposing goods market clearing, I get that in equilibrium, wages satisfy

$$w_i L_i^W = \sum_{n=1}^{2456}\pi_{ni}Y_n^R = \sum_{n=1}^{2456}\frac{L_i^W A_i^{\sigma-1}(\delta_{ni}w_i)^{1-\sigma}}{\sum_{j=1}^{2456}L_j^W A_j^{\sigma-1}\left(\delta_{nj}w_j\right)^{1-\sigma}}Y_n^R \tag{19}$$

By Walras Law, this also means labor market clearing. So, just as in Pérez-Cervantes (2013), it is possible to obtain the vector of productivities by iterating over equation 19 using the data. Defining the source effect of municipality i as $S_i = L_i^W A_i^{\sigma-1}(w_i)^{1-\sigma}$, and finally to obtain trade costs δ_{ni}, I follow the methodology in Pérez-Cervantes and Sandoval-Hernández

[26] This probability distribution has very simple properties, namely $\rho(i|n) = \frac{B_{in}w_i^\lambda}{\sum_{j=1}^{2456}B_{jn}w_j^\lambda} \geq 0$ which means $\sum_{n=1}^{2456}\rho(i|n) = 1$. This distribution is exogenous in this paper, and I won't study the reasons why this distribution might change over time. In other words, B_{in} will be taken as given and unobserved.

(2015) making use of the *Red Nacional de Caminos*, also recently published,[27] where

$$\delta_{ij} = \begin{cases} e^{0.0557+0.0024d_{ij}} & i \neq j \\ 1 & i = j \end{cases}$$

and d_{ij} is the number of hours that it takes to go from municipality j to municipality i using the optimal route, I get that this equation becomes

$$Y_i^W = \sum_{n=1}^{2456} \frac{S_i \left(\delta_{ni}\right)^{1-\sigma}}{\sum_{j=1}^{2456} S_j \left(\delta_{nj}\right)^{1-\sigma}} Y_n^R \tag{20}$$

where $Y_i^W = w_i L_i^W$ is the total wages paid in municipality i.

2.5 Real income and welfare

Real income in municipality n is the wages of the workers who live there deflated by the local price index. Note that the price index has also a simple expression given the notation for the source effect

$$p_n = \left(\sum_{j=1}^{2456} S_j \left(\delta_{ni}\right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{21}$$

which means that the real income is obtained up to a constant,[28] so I remove all the constants that cannot be obtained directly from the data because they are positive and so they preserve

[27]In the cited paper, it is shown that average transport costs for goods take this functional form for a large set of robustness checks and trade model specifications.

[28]It is the same constant in every municipality, and equals $\frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}F} \left(\overline{A}\right)^{\sigma-1}$ where $\overline{A}$ is the level of productivity that would equate the model's expenditure of a given municipality with the one in the data.

the rank, letting the expression of real income as

$$\sum_{i=1}^{2456} \frac{\rho(i|n)\, w_i}{p_n} = \frac{v_n}{p_n} = \frac{Y_n^R}{p_n L_n^R} \tag{22}$$

As for the municipal welfare, the friction that was imposed that consumers take their municipality of residence as given, leaves welfare impossible to identify, which is the main reason why in this paper I won't be comparing welfare among municipalities, or in other words, the main reason why I do not study migration.[29] However, it is important to note that by allowing commuting instead of migration in the model, the parameter restriction $\lambda > 1$, irrelevant for the analysis and reach of this paper, implies that the mean indirect utility of resident ϕ of municipality n working in i will be

$$\begin{aligned} \mathbb{E}\left(U_{in\phi}\right) &= \mathbb{E}\left(\frac{(B_{in})^{\frac{1}{\lambda}}\, b_{in}(\phi)\, w_i}{p_n} \,\middle|\, \frac{(B_{in})^{\frac{1}{\lambda}}\, b_{in}(\phi)\, w_i}{p_n} \geq \max_{j\neq i} \frac{(B_{jn})^{\frac{1}{\lambda}}\, b_{jn}(\phi)\, w_j}{p_n}\right) && (23) \\ &= \Gamma\left(\frac{\lambda-1}{\lambda}\right) \frac{1}{p_n} \left(\sum_{j=1}^{2456} B_{jn} w_j^{\lambda}\right)^{\frac{1}{\lambda}} && (24) \end{aligned}$$

where the term $\Gamma\left(\frac{\lambda-1}{\lambda}\right)$ is finite and positive only if $\lambda > 1$. Notice that the expected utility is increasing in the wages of any municipality, and decreasing in the price index of the municipality of residence. All the residents ϕ have the same expected utility, regardless of the realizations of the amenities random variable in the rest of the municipalities, and independent of the municipality where they work. This last fact will not be exploited in this

[29] See Allen and Arkolakis (2014) and Allen and Arkolakis (2016) for seminal contributions to modeling migration with gravity.

paper, but is the key ingredient to estimate the value of time for intermunicipal commuters in Mexico City in Pérez-Cervantes and Cuéllar (2016).[30] Welfare in municipality n, therefore, will not need to compute v_n, but this number will be important to obtain the real income, which is the main focus of study of section 4.

2.6 Changes over time

Following standard notation,[31] let x' be the value in 2015 of a variable that had a value of x in 2010. Then, the gross change in x, denoted as $\widehat{x}$ is x'/x. It is straightforward to show that the gross change in the price indices satisfies:

$$\widehat{p}_n^{1-\sigma} = \sum_{i=1}^{2456} \pi_{ni} \widehat{L}_i^W \left(\frac{\widehat{\delta}_{ni} \widehat{w}_i}{\widehat{A}_i} \right)^{1-\sigma} \tag{25}$$

Note two things. First, that the price index gross change can be less than one (i.e. there can be a reduction in the price index) due to the fact that there is an increase in the number of workers in Mexico (or even more primitively in the model, a reduction in the fixed cost parameter F, but this option will not be considered. This document assumes σ to be fixed over time). So it will be convenient to express the gross change in the number of workers in a municipality as the product of the gross change in municipal worker share ($\widehat{\ell}_i^W$) times the

[30] It is possible to obtain the normalized values of B_{in} for every i and every n using the commuting data, where every probability is interpreted as a share, invoking the Law of Large Numbers. As it is clear in equation 7, the constant of normalization is by columns of the matrix B_{in}, so $\sum_{j=1}^{2456} B_{jn} = 1$ without loss of generality. This is fine since it is assumed that the location of residence is fixed, so it is expected that average utility is equated within the municipality of residence, but impossible to compare across municipalities.

[31] See Dekle, Eaton, and Kortum (2007) for an extraordinary introduction to this notation.

gross change in the total workers in Mexico ($\widehat{L}$):

$$\widehat{L}_i^W = \widehat{\ell}_i^W \widehat{L} \tag{26}$$

The second thing to note is that since the source effect is normalized, it means that productivity is normalized too. What I do for normalization is to divide the country's population in two halves. One half will consist of municipalities that are below the median productivity (normalized to be equal to 1) and the other half will have municipalities with above median productivity. So, it is convenient, again, to express the gross change in productivity in a municipality as the product of the gross change versus the median ($\widehat{a}_i$) times the gross change in the productivity of the median municipality ($\widehat{A}$):[32]

$$\widehat{A}_i = \widehat{a}_i \widehat{A} \tag{27}$$

which means that the change in price index is:

$$\widehat{p}_n = \frac{\widehat{L}^{\frac{1}{1-\sigma}}}{\widehat{A}} \left(\sum_{i=1}^{2456} \pi_{ni} \widehat{\ell}_i^W \widehat{a}_i^{\sigma-1} \left(\widehat{\delta}_{ni} \widehat{w}_i \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{28}$$

The gross change in resident income equals

$$\widehat{v}_n = \sum_{i=1}^{2456} \frac{\rho\left(i|n\right) w_i}{v_n} \widehat{\rho}\left(i|n\right) \widehat{w}_i = \sum_{i=1}^{2456} \beta_{in} \widehat{\rho}\left(i|n\right) \widehat{w}_i \tag{29}$$

[32]The median is an arbitrary statistic, and this procedure works as long as productivity is measured relative to the same municipality in both periods of time.

where β_{in} is the percentage of the resident income in municipality n that comes from its own residents who work in municipality i. In financial terms, this is a measure of the exposure that municipality n has on the income sourced from i. So, given a reduction in their own wages, municipalities that initially had smaller exposure are hit the least, so as municipalities that can adjust the probability $\widehat{\rho}\left(n|n\right)$ to the lowest possible number.

As for real income, its gross change is

$$\frac{\widehat{v}_n}{\widehat{p}_n} = \frac{\widehat{A}\widehat{L}^{\frac{1}{\sigma-1}} \sum_{i=1}^{2456} \beta_{in}\widehat{\rho}\left(i|n\right)\widehat{w}_i}{\left(\sum_{j=1}^{2456} \pi_{nj}\widehat{\ell}_j^W \widehat{a}_j^{\sigma-1}\left(\widehat{\delta}_{nj}\widehat{w}_j\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}} \tag{30}$$

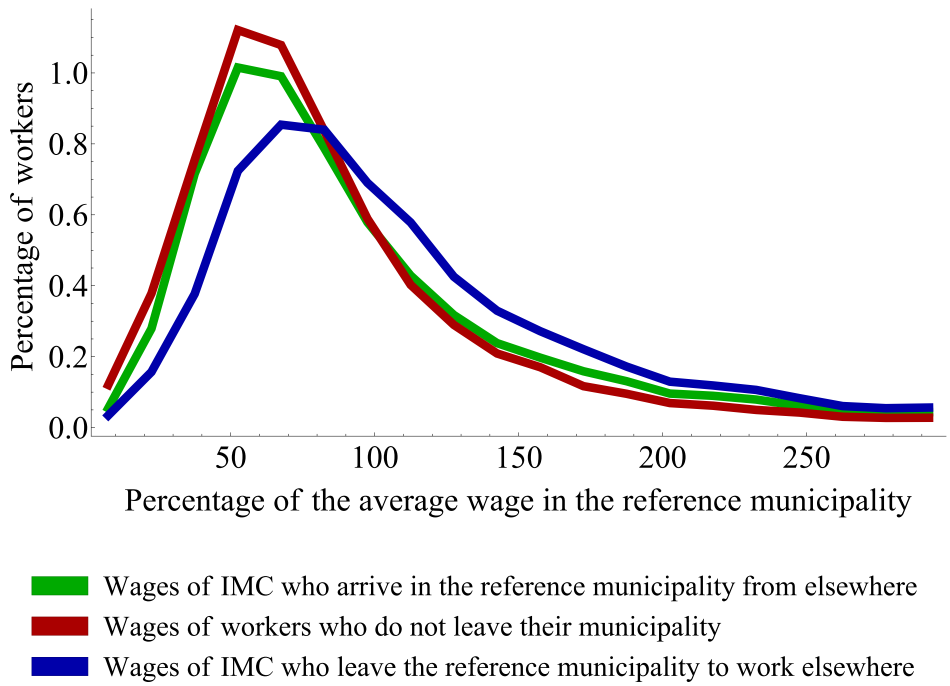
which can be interpreted as the ratio of two exposures: the one coming from income (β_{in} can be interpreted as the probability that \$1 of municipal income in n came from municipality i) and the one coming from expenditure (π_{ni} can be interpreted as the probability that \$1 was spent in goods from municipality i given that \$1 was spent in municipality n).

3 Data

I have data from the recently published *Muestra Especial del Censo de Población y Vivienda 2010* (2 million homes) and the *Encuesta Intercensal 2015* (6 million homes), where the questions answered include the location of residence, the location of work, the salary, and the mode of transport.[33] I use this data to obtain average wages, total resident wages and

[33] Commuting time is also reported in 2015. It is not punctually reported but instead in 30 minute intervals. I discard this extremely valuable information, because the non-pecuniary aspects of commuting are not studied here, and to be able to do comparisons with 2010 without having to assume anything about commuting time in

Figure 1: Distribution of the wages for commuters and non-commuters



Source: Calculated by the author with data from *Encuesta Intercensal 2015* by INEGI.

total worker wages at the municipal level. The data for 2015 show that IMC gain an average 30% higher salary compared to the workers in the same municipality who do not commute. Analogous calculations show that workers who arrive to a municipality get not only almost the same average wages, but also almost identical distribution of wages (see figure 1). That is, wages (and productivity) seem inherent to the municipality, not to the workers.[34]

2010.

[34]This last fact can be explained by the other factors of production, such as capital, skill distribution demands, and even land and weather. These factors are take as given from the point of view of the worker. So, in any case, those combinations of factors attract the same types of workers both locally and from other municipalities.

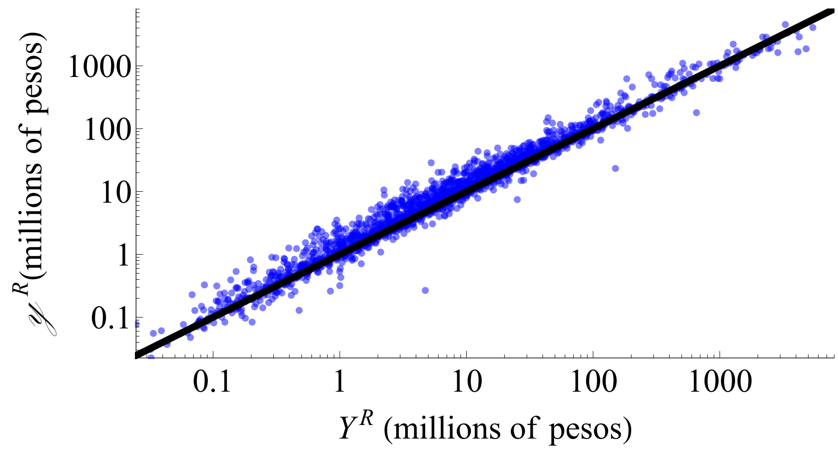
3.1 Constructing wages and income from the data

Taking this model to the data is relatively easy, but first I will clarify the construction of the municipal income, from the data. Section 4 has both the results of all the ways the variable construction is performed. It is straightforward to get the values of L_i^W, L_i^R, and Y_i^W just by counting the number of people who answered working in a municipality, living in a municipality, and the total income of the people who answered working in the same municipality. The wages in every municipality will be the average wage paid to the workers of the same municipality, irrespective of where they live

$$w_i = \frac{Y_i^W}{L_i^W} \tag{31}$$

Then, there are two possibilities for constructing municipal income. The first one is just getting $\rho(i|n)$ for every pair of municipalities by counting the number of people who answered living in n and working in i, and dividing this number by L_n^R. Then, municipal income is just as defined in equation 18. The other one is just adding the income of the workers who answered that they live in municipality n and call this number $\mathscr{Y}_n^R$. As it can be seen in figure 2, there was not much difference in the total income by municipality in 2010. The same holds for 2015. It is noticeable that the predicted mean income (obtained with the commuting probabilities) is higher for middle-income municipalities. This might imply the presence of self-selection in the decision of commuting, something that will not be studied here. It follows that equation 20 needs a municipal income for the right hand side, that is,

Figure 2: Total Income by the workers residing in a municipality, and total expected income if every commuter to the working municipality got the average wage of that municipality



Source: Calculated by the author with data from *Muestra Especial del Censo de Población y Vivienda 2010* by INEGI.

either $\mathscr{Y}_n^R$ or Y_n^R. Noticing that $\sum_{i=1}^{2456} Y_i^W = \sum_{i=1}^{2456} Y_i^R = \sum_{i=1}^{2456} \mathscr{Y}_i^R$ it is possible to define residents' income as the workers' income plus some deficit. That is,

$$Y_i^W = \sum_{n=1}^{2456} \frac{S_i (\delta_{ni})^{1-\sigma}}{\sum_{j=1}^{2456} S_j (\delta_{ni})^{1-\sigma}} \left(Y_n^W + D_n\right) \tag{32}$$

where either $D_n = Y_i^R - Y_i^W$ or $\mathscr{Y}_i^R - Y_i^W$. The last chapter of Pérez-Cervantes (2013) shows (in a very different context and answering a very different question) that this system has a unique solution for the vector of S regardless of the deficits per municipality, as long as $\sum_{i=1}^{2456} D_i = 0$. This means that, as long as labor income can be labeled to the municipal workers (and not to the residents, just as almost all the surveys do), this model specification works to obtain productivities. The same chapter also shows that assigning the income to the residents and giving a negative deficit to the workers' payments leads to incorrect solutions for the source effect vector. In other words, the income in the right hand side of equation 32 is irrelevant for obtaining productivities, as long as Y_i^W is not replaced by Y_i^R on the left hand side of the same equation.

This paper will use both Y_n^R and $\mathscr{Y}_n^R$ as municipal income, and note the qualitative differences in the results, if any. With $\sigma = 4.5$ (as in Monte, Redding, and Rossi-Hansberg (2015)), the fixed point for S is found and the municipal productivity is just

$$A_i = w_i \left(\frac{S_i}{L_i^W} \right)^{\frac{1}{\sigma - 1}} \tag{33}$$

Note that if any other type of model in the EK literature was used, the productivity would have to be $\left(L_i^W\right)^{\frac{1}{\sigma-1}}$ times larger. Given the huge variations in municipal working population, this model helps to flatten the productivity parameters into a much more reasonable average marginal cost distribution at the municipal level. Finally, figure 3 shows that there are hundreds of municipalities in Mexico that source more than half of their wages from other municipalities (small β_{nn} when the income notation is Y_n^R), and also that most municipalities have an extremely large loading of self-exposure (large β_{nn}), that is, most of their income is obtained from non-commuters. That is, there is large variation in exposure to local wage shocks, meaning that it is likely that workers will react differently to a local wage shock, mostly depending on the exposure to their local wages.

Figure 3: Distribution of Self-exposure of income in 2015

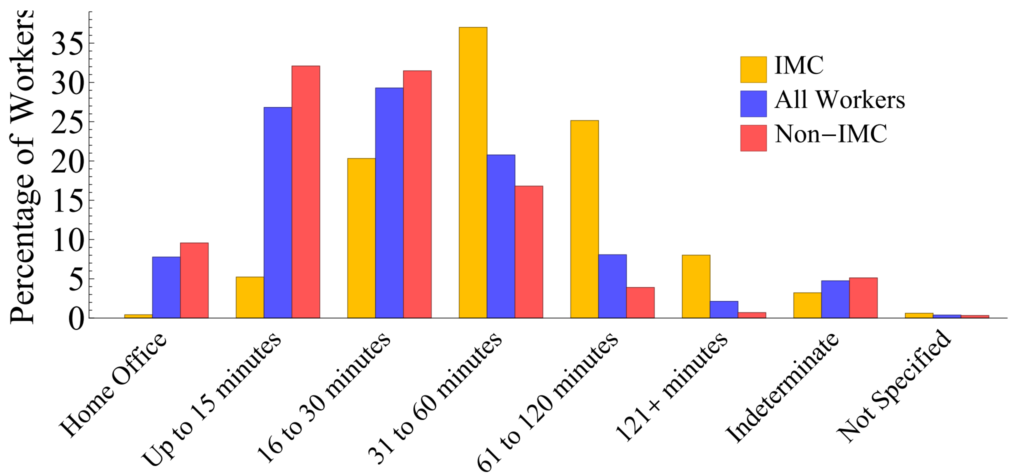
Source: Calculated by the author with data from *Encuesta Intercensal 2015* by INEGI.

3.2 Summary statistics of the commute

In this subsection I am writing some summary statistics about the commutes of the Mexican workers, although it won't be of any use in the analysis of this document. This is done with the only purpose of giving the reader a clearer panorama of the dimensions of the return of the commute in Mexico, which as I just showed, implies a 30 percent average gain for the ones who do the investment of commuting, but so far I have not mentioned the costs. This subsection will give an idea of what can be thought of as the cost of commuting by showing some statistics about the commuters who actually incur in the cost of traveling to another municipality every day.

In the year 2015, which is the only year with self-reported commuting time in the data set, the workers who responded being IMC traveled a median of between 31 and 60 minutes, as pictured in figure 4, where it can also be seen that almost 75 percent of non-IMC workers

Figure 4: Self-reported commuting time for IMC in 2015



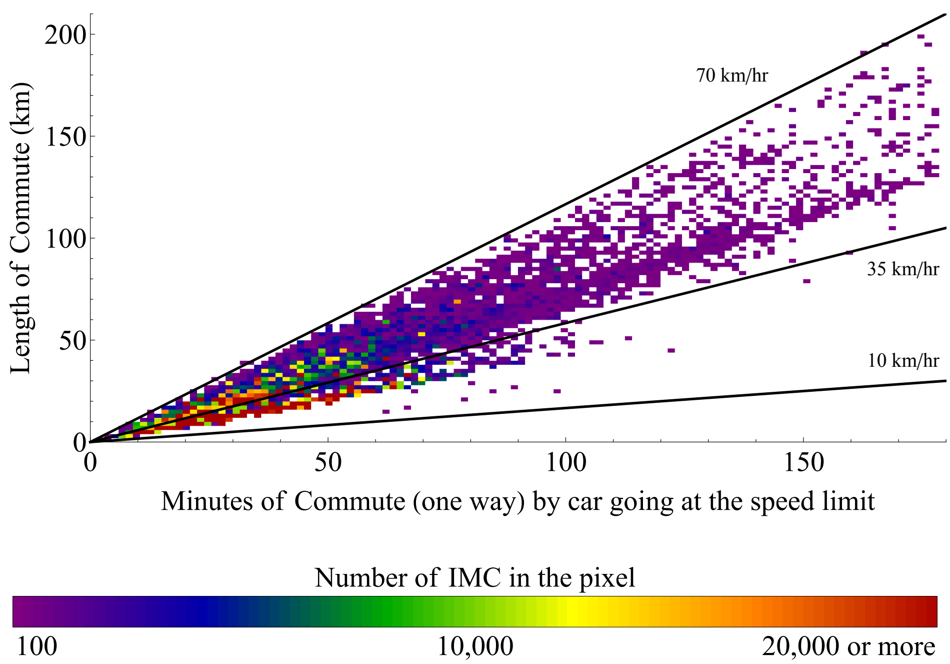
Source: Calculated by the author with data from *Encuesta Intercensal 2015* by INEGI.

take less than an hour round-trip to work. Not surprisingly, IMC have a longer commute, and using the mean of the reported bins as the exact commuting time, they spend on average more than 6 hours more per week in the commute.

Based on the location of their residence and of their place of work, I have calculated, using the *Red Nacional de Caminos* data, the length of the fastest possible commute for every worker in the sample, that is, a total of more than 80 million individual queries, counting both 2010 and 2015.[35] The results for 2015 are in figure 5, where it is possible to see that most workers go to locations that are *potentially* fast to arrive by car if it were going at the speed limit, and not necessarily close in distance. If the real commute of the IMC took the same path but at a slower speed, the results have the pixels biased to the left. And since the commute cannot take a shorter time but can in fact be vary in distance (taking a *shorter* path with a *more* than proportionally *lower* average speed, or taking a *longer* path with a *less* than proportionally

[35] See Pérez-Cervantes and Cuéllar (2016) for a fully detailed description on how to calculate each pair of paths, lengths, and travel times.

Figure 5: Population Distribution of the minutes of commute and of the length of the commute for IMC, 2015.

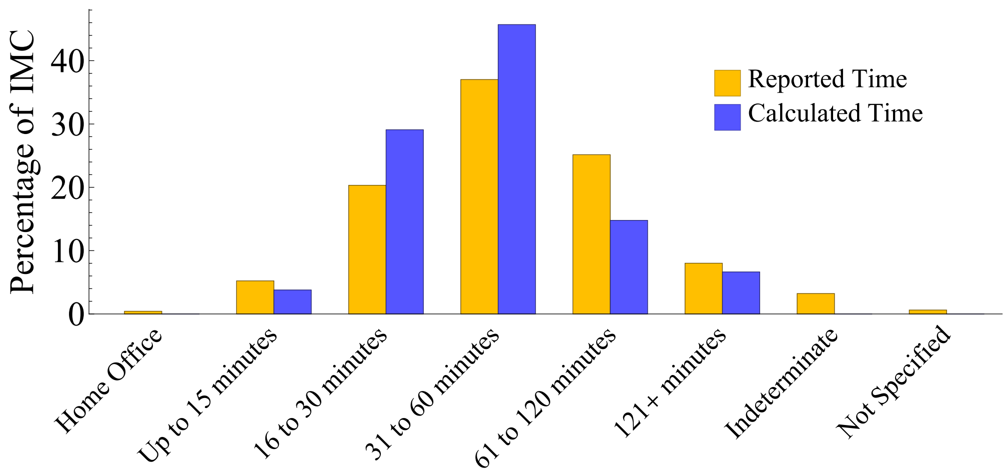


Source: Calculated by the author with data from *Encuesta Intercensal 2015 and Red Nacional de Caminos* by INEGI.
Note: Pairs of time and distance that had less than 100 workers are not on the plot.

higher average speed) it is not direct to conclude that the *real* but unobserved colored pixels representing the commuting routes are above or below the pixels of figure 5, but for sure to the right. Comparing the calculated time with the self-reported time, I get figure 6, where it is shows that the difference between the calculated and the real is only a few minutes. It also shows that the calculated distribution is downward biased on the left of the distribution and upward biased in the right tail of the distribution, which is consistent with the calculated commutes being a little biased towards the left part of figure 5.

The final ingredient about the commuting, the municipal pair amenity parameter B_{ni} that contains characteristics that are specific of the pair of municipalities of residence n and of

Figure 6: Distribution of the time length of IMC commute (one way). Reported and calculated, 2015.



Source: Calculated by the author with data from *Encuesta Intercensal 2015 and Red Nacional de Caminos* by INEGI.
Note: Calculated times do not have the categories of home office, indeterminate and not specified.

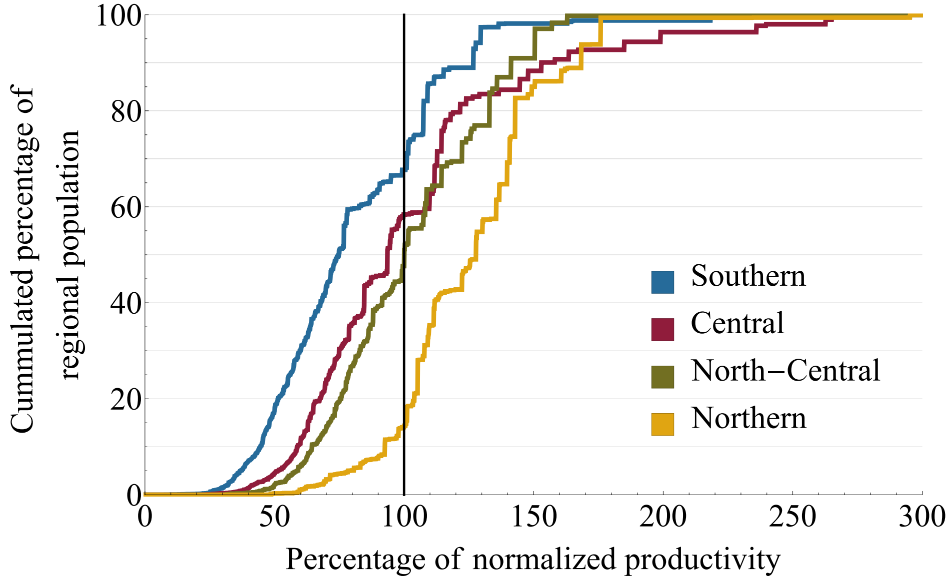
work i such as the distance and or cost of the commute, the weather of the workplace, etc. will be estimated for every i for both 2010 and 2015 as residuals, but neither the magnitudes nor the changes over time will be studied.

4 Results

The first thing that was done was to obtain the productivities of the municipalities iterating over equation 20. Municipal income is $\mathscr{Y}_i^R$, extracted directly from the data. Productivity by region of Mexico in 2015 is pictured in figure 7.[36] There, it is possible to see that most of the

[36]Northern region: Baja California, Chihuahua, Coahuila, Nuevo León, Sonora, and Tamaulipas. North-Central region: Aguascalientes, Baja California Sur, Colima, Durango, Jalisco, Michoacán, Nayarit, San Luis Potosí, Sinaloa, and Zacatecas. Central region: Ciudad de México, Estado de México, Guanajuato, Hidalgo, Morelos, Puebla, Querétaro, and Tlaxcala. Southern region: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz, and Yucatán. See Banco de México (2014) or any other regional economies report to obtain the regional characteristics of this geographic classification.

Figure 7: Distribution of productivity of every municipality in Mexico



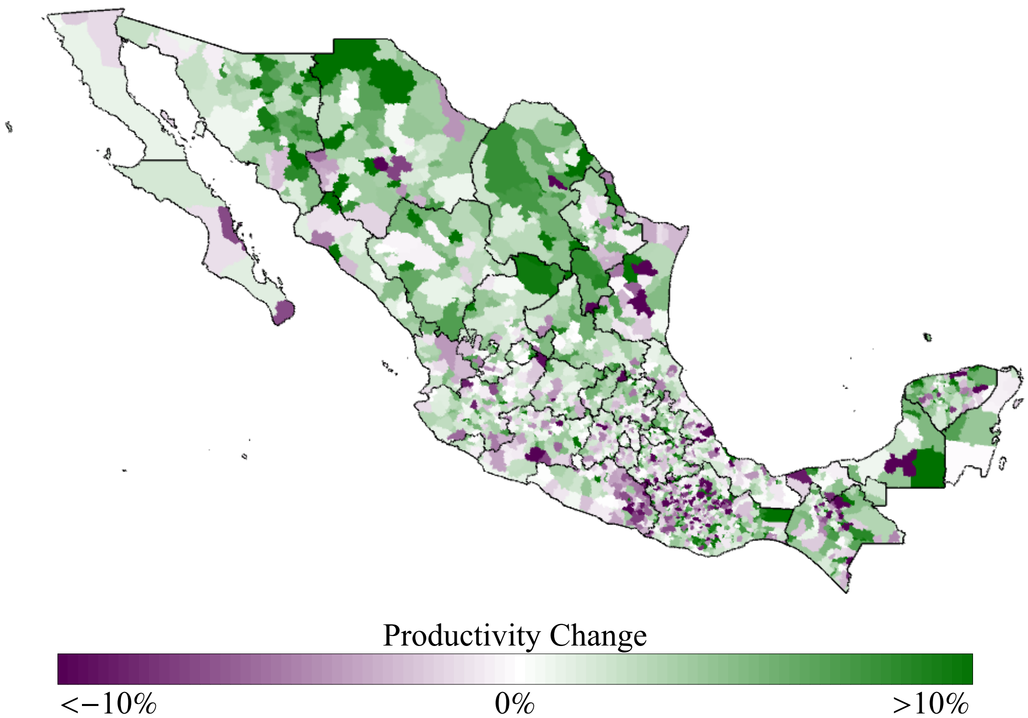
Source: Calculated by the author with data from *Encuesta Intercensal 2015 and Red Nacional de Caminos* by INEGI.

population in the Northern region lives in productive municipalities, and that the region with the most people living in high productivity municipalities is the Central region. Then, I get the change in productivity assuming $\widehat{\delta}_{in}$ to equal the change in transport costs derived from the construction of the Tuxpan-Mexico City and Durango-Mazatlan highways, for all i and all n.[37] The changes in productivity are pictured in figure 8, where it is clear that productivity shocks have a positive spatial correlation.

Finally, I plot the change in real income of equation 30 assuming $\widehat{\widehat{AL}}^{\frac{1}{\sigma-1}} = 1$ (the qualitative results are invariant to this assumption) against the change in productivity of figure 8. Again, municipal income is $\mathscr{Y}_i^R$. There, it is evident that the municipalities that suffered a negative shock adjusted to it (recall from figure 3 that most municipalities have an overwhelming

[37]See Pérez-Cervantes and Sandoval-Hernández (2015) for an extensive discussion on how are these gross changes in transport costs between these two years obtained, and why the relevant cost is still the cost of time.

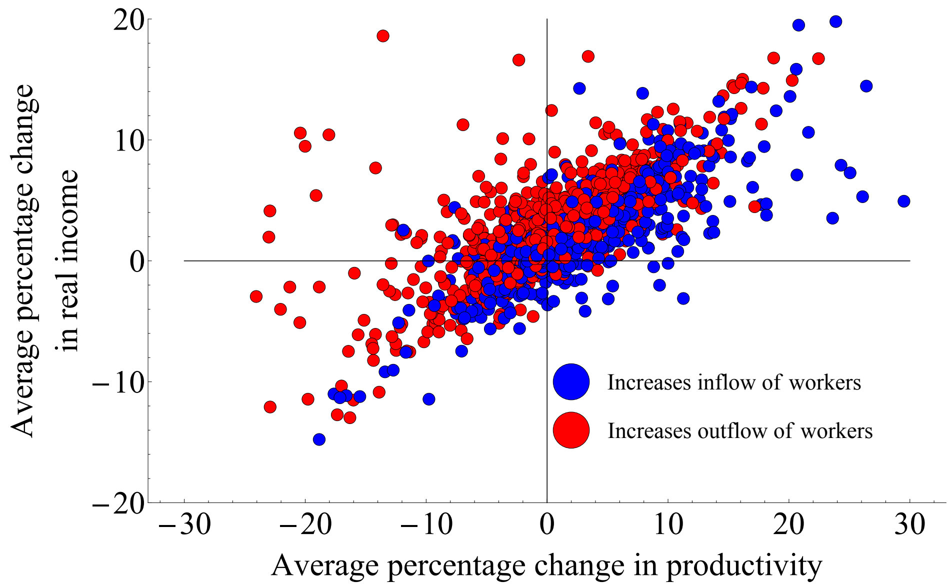
Figure 8: Changes in productivity between 2010 and 2015



Source: Calculated by the author with data from *Muestra Especial del Censo de Población y Vivienda 2010, Encuesta Intercensal 2015, and Red Nacional de Caminos* by INEGI.

share of their income coming from non-commuters) by increasing the number of workers that commute. In fact, there are some municipalities that saw an increase in their real income. Most of those municipalities increased the number of intermunicipal commuters. This positive real effect is an interaction between the lower price index faced from the reduction in wages and a relatively good adjustment in the number of workers going to other municipalities that did not suffer a negative productivity shock. On the other hand, municipalities that saw an increase in productivity also saw an increase in the number of workers from other municipalities that decided to go to work there. Some of these workers were even residents of the municipality who worked elsewhere before. I repeat the exercise

Figure 9: Changes in productivity and changes in real income between 2010 and 2015



Source: Calculated by the author with data from *Muestra Especial del Censo de Población y Vivienda 2010, Encuesta Intercensal 2015, and Red Nacional de Caminos* by INEGI.

with municipal income equal to Y_i^R and the results are qualitatively identical.

5 Conclusion

This paper is able to identify that intermunicipal commuters travel for higher wages (although commuters and non-commuters all have the same expected welfare). On average, intermunicipal commuters earn 30 percent more than their non-commuting counterparts. There is a large variation in the proportion of municipal income that comes from commuters, but most of the municipalities have a large share of their income non-commuters.

Most of what we can learn about the dynamics of commuting in this paper is in equation 30. In order to have sustained growth, the necessary infrastructure is needed to be built to increase productivity (large $\widehat{A}$), but if this is not possible in every municipality, a transportation infrastructure that allows a larger number of intermunicipal commuters to municipalities with large productivity is also desirable (small $\widehat{\rho}\,(i|n)$ wherever $\widehat{w}_i$ is small, large $\widehat{\rho}\,(i|n)$ wherever $\widehat{w}_i$ is large).

Finally, there is evidence that workers change the location of their workplace based on the returns from commuting. In particular, between the years 2010 and 2015 in Mexico, the average worker living in municipalities that suffered a reduction in productivity smoothed the shock by picking another location where to work. There were some municipalities, however, that could not absorb the shock with commute and saw a reduction in their real wage. Analogously, municipalities that saw an increase in productivity also saw an increase in the number of workers who decide to go there to work, and this increase in workforce lowered their price indices, reinforcing the positive shock.

References

ALLEN, T., AND C. ARKOLAKIS (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129(3), 1085–1140.

——— (2016): "The Welfare Effects of Transportation Infrastructure Improvements," .

BANCO DE MÉXICO (2014): "Reporte sobre las economías regionales Enero-Marzo 2014," *Banco de México*.

DEKLE, R., J. EATON, AND S. KORTUM (2007): "Unbalanced Trade," *The American economic review*, 97(2), 351–355.

EATON, J., AND S. S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

HELPMAN, E. (1998): "The size of regions," *Topics in public economics*, pp. 33–54.

KRUGMAN, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–499.

MONTE, F., S. J. REDDING, AND E. ROSSI-HANSBERG (2015): "Commuting, migration and local employment elasticities," *NBER Working Paper*, (w21706).

PÉREZ-CERVANTES, F. (2013): "Railroads and Economic Growth: A Trade Policy Approach," *Thesis (Ph.D.) – The University of Chicago, Division of the Social Sciences, Department of Economics*.

PÉREZ-CERVANTES, F., AND Ó. CUÉLLAR (2016): "Estimating the Value of Time in Mexico City," *Forthcoming. Documento de Investigación Banco de México*.

PÉREZ-CERVANTES, F., AND A. SANDOVAL-HERNÁNDEZ (2015): "Estimating the Short-Run Effect on Market-Access of the Construction of Better Transportation Infrastructure in Mexico," *Documento de Investigación 2015-15. Banco de México*.