Trade, Workers and Cities*

El-Mehdi Aboulkacem† Clément Nedoncelle ‡

November 26, 2018

Preliminary version. Please do not quote or circulate.

Abstract

We estimate the causal impact of international trade shocks on spatial allocation of workers. We test whether trade shocks lead workers to live further away from their working place. We provide a theoretical framework emphasizing the role of trade shocks on residential choices of workers. We find empirical support for the micro mechanism using employer-employee data for the French *Ile-de-France* region (surrounding Paris). We estimate that increases in wages coming from exogenous trade activities increase commuting distance for these workers. Following these shocks, workers tend to move towards richer and denser cities. Our results cast novel insights upon the consequences of trade activities on spatial allocation of workers.

Keywords: Distance; Firms; Trade; Labor Market

JEL Codes: F16, F61, R1, R40

*This paper previously circulated under the title "Trade Shocks, Spatial Sorting and Urban Sprawl". This research did not benefit any financial support. We also would like to thank Hubert Jayet, and participants at the Canadian Economics Association Meeting in Antigonish (2017), at the ETSG 2018 as well as seminar participants in Lille and INRA.

†IFSTTAR - SPLOTT; Email: el-mehdi.aboulkacem@ifsttar.fr

‡Corresponding Author. INRA-0210 Economie Publique, Université Paris-Saclay; Email: clement.nedoncelle@inra.fr

1 Introduction

Do workers move because of trade shocks? While the consequences of trade activities on firms outcomes have been extensively documented both at the micro and at the aggregate level, studies focusing on the effect of exporting activities on other worker-level outcomes remain limited both in number and in scope. Most of the existing literature focuses on the wage effect of trade (for instance Card et al. (2018)) but remains silent regarding the consequences of trade on other outcomes through wages. The present paper fills this gap by investigating the consequences of trade activities (identified at the firm level) on the location decision of workers, that is how far they live from their working place.

A large set of trade theories helps to understand the impact of export activities across workers' wages. Standard trade theories predict that wage across occupations are differently affected by trade expansion. New trade theories, based upon firm heterogeneity, reach similar conclusions. A clearcut conclusion from this literature supports that trade activities generate changes in wages, across and within groups. Consequences of these changes in wages have however not been studied yet, in particular regarding the location decision of workers. On the other hand, residential and location decisions of workers are important choices made by households and have non-trivial consequences on structure and organization of cities. In a broad view, city boundaries are shaped by households (and firms) locations. Understanding the determinants of housing location is thus of particular importance. This is all the more important as trade activities are expanding, and this for many reasons: fall of transportation costs, increase in technology choices and political decisions all shape trade patterns at the aggregate and at the firm level.

In the present paper, we study the impact of trade activities at the firm level on location decision of workers directly affected by trade activities. We start by providing a stylised theoretical model capturing the spatial sorting of workers following changes in wages. We consider a monocentric city, with two kinds of firms –exporters and non-exporters - and two kinds of workers – for exporters and for non-exporters–. Trade shocks straightforwardly affect exporters employees wages while domestic firms's wages are not affected. The partial equilibrium effect is such that, at the individual exporter employee level, wage increases leads to increase in their bid-rent, and to increase in commuting distance.

We provide micro-level support for this result. We estimate the worker-level commuting distance elasticity to wage, using data covering the universe of Ile-de-France (surrounding Paris). We use commuting times data between 1600000 couples of origins-destinations. This unique dataset is matched with an exhaustive employer-employee dataset providing information about workers' earnings, living and working cities. We are confident in our analysis to provide estimates of the causal impact of trade on location. We identify trade shocks that are exogenous to firms and workers by computing, using disaggregated customs data, the world import demand that is addressed to the firm, in the spirit of Mayer et al. (2016). We instrument individual wages with this measure of trade shocks in an IV-2SLS estimation of the distance-wage elasticity. We hence identify the effect that is arising from trade shocks only and not by other simultaneous shocks to firms or workers that may be related to locations.

We estimate that workers that experience higher wage, because of trade shocks, also live farther away from their workplace, ceteris paribus. Our fixed effect structure ensures that this effect occurs for any worker across years: as a benchmark estimate, we find that a 1% increase in wage raises commuting distance by 0.3% on average. Due to data constraints, our estimates are short-run estimates and are hence small compared to existing estimates. We identify short-run changes in location following changes in wages and are thus a low benchmark for the actual elasticity. Our results resist to a set of robustness checks on alternative mechanisms, subsamples and measures of key variables.

Related Literature and Contributions Our paper is related to different streams of literature. The first is the literature on an income-commuting relationship. Most of the existing theoretical work in urban economics on this topic follows the Alonso (1964)- Muth (1969) basic urban economic model that supports a positive income-commuting relationship: those with higher incomes have higher commuting distances. Additional assumptions – such as including endogeneous house prices (Straszheim, 1987) or moving costs (Zenou, 2009) – all confirm this positive relationship. There are also several alternative labour market theoretical explanations for a relationship between income and commuting distance. First, in both competitive and imperfect labor market framework, firms located far from residences compensate their workers with appropriately higher wages, which implies a spatial wage gradient (e.g., Fujita et al., 1997, Mulalic et al., 2014, Timothy and Wheaton, 2001). On top of this, Zenou (2009) supports that wages are set independent of the length of the commute in a standard monocentric urban model where employers have monopsony power, because house prices fully compensate for the length of the commute.

On the empirical side, existing estimates are consistent with both urban and labor economic theory, but remain mainly silent regarding the causal impact from income to commuting and location decisions. Some papers, such as Gutierrez-i Puigarnau et al. (2016), acknowledge both the existence of unobserved factors and reverse causality issues, deal with this issues by using sample restrictions: they argue that reverse causation is less likely to be at play when workers move residence but do not change employer and thus estimate the elasticity on this sample of workers. As a counterpart, their result can solely be interpreted as a long-run elasticity. As an alternative to using sample restrictions, a standard method for dealing with these endogeneity difficulties is IV estimation. The main problem with this approach is to find suitable instruments for the income variable. Candidates are variables affecting income but not locations. Our paper provides a robust and convincing instrument for individual wage: we identify trade shocks that are exogenous to firms and workers decisions. We thus contribute to this literature by providing a robust and causal estimate of the distance elasticity to wage.

Second, our analysis also builds on the literature on trade with heterogeneous workers. While traditional theories address causes and consequences of trade using aggregate factor models, a recent literature introducing worker heterogeneity into models has emerged. This trends echoes the parallel increased availability of employer-employee datasets. Understanding how heterogeneous workers sort into industries and match with heterogeneous firms (in productivity, mainly) provides additional insights on the effect of trade opening on wage distribution and on factor allocation. Extensive reviews of this literature are provided in Davidson and Sly (2012); Grossman (2013). Not only does liberalized trade favor the matching between workers and firms (Davidson et al., 2012), but exporters and non-exporters match with sets of workers that are different (Bombardini et al., 2015). With respect to this literature, we bring a particular attention to spatial aspects of worker heterogeneity: our paper rationalizes a consequence of spatial heterogeneity in worker's type and how trade affects the spatial distribution of workers.

The paper is structured as follows. The next section provides the model and captures the effect of trade on wages and on spatial allocation of workers. Section 3 presents the main dataset we use in the analysis and Section 4 describes the empirical strategy. Section 5 provides the main results. The final section concludes.

2 Theoretical model

2.1 Set-up

We consider an open monocentric city with a transport system that is radial, dense, free of congestion and only used for home-workplace trips. In this city, land is a featureless plain where all the parcels are identical and ready for residential use. In this context, the only characteristic of interest of each location within the city is its distance from the center. Thus, the urban space can be treated as a ray starting from the center of the city.

Two kinds of firms are present in the city : exporting firms producing varieties of a good sold on the world market and domestic firms producing varieties of a good exclusively intended for the local market. Firms compete in a standard monopolistic competition such that the producer of each variety has a monopoly position in its sub-market.

Two categories of workers reside in the city : workers intended to domestic firms and workers to the firms producing the good for the global market (called the global good in the rest of the paper). Workers cannot move from a domestic firm to and exporter and vice versa. Thus, their are two "independent" labor markets. Workers enjoy a perfect geographic mobility : on the one hand they can freely move within the city and on the other, they can leave it if they can enjoy a higher utility level elsewhere. Thus, at the equilibrium, workers living in the city have no incentive to move.

Finally, whatever the production sector (domestic / global good), firms can enter it freely, without any administrative, price or technology barriers.

2.2 Demand and wages

All workers have one unit of labor that they give inelastically supply to firms.

Domestic firms behavior. All domestic firms are identical. They all have an increasing returns to scale technology, with a fixed production cost corresponding to the amount of labor necessary to initiate the production, and a marginal constant cost corresponding to the remuneration of the work necessary for the effective production once it is initiated. The domestic firms fixed cost is small and cannot be shared between them. Thus, the quantity of labor needed to produce $Q_{d,j}$ unit of the j^{th} variety of the domestic good is[1] :

$$l_{d,j} = f_d + n_d Q_{d,j}$$

where f_d is the quantity of labor needed to initiate the production, and n_d the quantity of labor needed to produce an extra unit ans thus an inverse measure of firm efficiency.

Let w_d denote the average wage in domestic firms. As each worker has one unit of labor, the production cost of $Q_{d,j}$ units is

$$C\left(Q_{d,j}\right) = w_d l_{d,j} = w_d\left(f_d + n_d Q_{d,j}\right)$$

Thus, in accordance with firms free entry and monopoly pricing rules, we obtain the standard result

$$p_{d,j} Q_{d,j} = w_d\left(f_d + n_d Q_{d,j}\right) = \frac{\sigma_d}{\sigma_d - 1} n_d w_d \tag{1}$$

where $p_{d,j}$ is the unit price of variety j and $\sigma_d > 1$ is an elasticity of substitution term reflecting the city inhabitants taste for diversity.

[1] This relation gives also the number of workers needed to produce $Q_{d,j}$ units.

It is worth noting that relation (1) states that if the demand for variety j increases and $p_{d,j}$, σ_d and n_d remain constant, the average wage of all the domestic firms workers increases at the equilibrium.

Exporting firms behavior. All the exporting firms have the same increasing returns to scale technology. The face a high fixed production cost that they can share if they agglomerate. Thereby, at the equilibrium, they agglomerate all in a same production cite and form a CBD that none of them has interest in leaving. The quantity of labor needed to produce $Q_{g,j}$ unit of the j^{th} variety of the exporting good can be written as follows :

$$l_{g,j} = f_g(N) + n_g Q_{g,j}$$

where N the number of firms agglomerating in the same production site, $f_g(N)$ the quantity of labor needed to initiate the production (decreasing in N) and n_g the quantity of labor needed to produce an extra unit.

If the average wage in the exporting firms is w_g, the cost for producing $Q_{g,j}$ unit of the global good j^{th} variety is :

$$(Q_{g,j}) = w_g l_{g,j} = w_g \left(f_g(N) + n_g Q_{g,j}\right).$$

The global good market is , in a long-term equilibrium. Furthermore, all its varieties are transportable without charges. Consequently, they are all bought at the same price p_g, considered by all the producers and consumers as taken.

If we note $\bar{N}$ the number of all the global good producers and N their number in the city $\left(\overline{N} \gg N\right)$, the unit price of the j^{th} variety is

$$p_{j,g} = \overline{N}^{\frac{1}{\sigma_g - 1}} p_g = \frac{\sigma_g}{\sigma_g - 1} n_g w_g$$

where $\sigma_g$the elasticity of substitution for the global good varieties.

Thus the average wage in the exporting firms is :

$$w_g = \overline{N}^{\frac{1}{\sigma_g - 1}} \frac{p_g}{n_g} \left(1 - \frac{1}{\sigma_g}\right) \tag{2}$$

If the international demand for the global good increases, it is very likely that $\bar{N}$ increases to. Thus, according to relation 2, if international demand for the global good increases and p_g, n_g and σ_g remain constant, the average wage of all the exporting firms workers increases at the equilibrium.

2.3 Wages and land use

As stated before, all the exporting firms agglomerate in the CBD. Concerning domestic firms, they have no incentive to do so. Moreover, we assume that they can move costlessly. Thus, at the equilibrium, on the one hand they locate far enough from each others[2] so their workers don't compete for the same land [3], and on the other they locate close enough from each other to let their workers ensure the residential continuity of the city. This continuity is supposed necessary for them because it allows a better diffusion of the varieties they produce in all the city.

Land is owned by absentee landlords, firms don't consume land, workers are risk neutral, indefinitely lived and can move costlessly.

The average wage in the exporting firms is w_g and in the domestic ones is w_d. However, within a firm, all workers are not all paid the same. In each sector, there is a base wage : Y_g in the exporting firms and Y_d in the domestic ones. The wage given to each worker depends on his efficiency, noted e_g in the exporting firms and e_d in the domestic ones where e_g and e_d are random variables higher than 1 and respectively lower than $E_g^{max}$and E_d^{max}. In this context, the wage of an exporting firm worker is e_gY_g and the one of a domestic firm worker is e_dY_d.

We suppose that the international demand for the global good is much higher than the one for the domestic good. Thus, according to 1 and 2, we can state that $Y_g > Y_d$ and $E_g^{max} > E_d^{max}$.

All the workers derive their utility from the consumption of housing and all the market goods varieties. We group here all these varieties in one market good, we note it z and we consider it as numeraire. If we note s the worker's house surface, the utility function describing his preference can be written as : $U = U(s, z)$.

The house surface and the quantity consumed of the market good depend on the location of area where the house is located. This comes from the fact that workers must commute from their homes to their workplaces. These trips entail monetary expenditures whose consequence is that the unit price of housing depends on the location in the city. As it is higher as we get closer from employment centers (Demonstration is coming later in the paper), worker must conduct a trade-off between living in a large house but far from their workcity and being close to there city but living in a small house.

We note $T_{g,0}(r) = \tau r$ the commuting cost of the workers employed by the exporting firms and $T_{d,i}(r) = \tau|r - d_i|$ the one of the workers employed by the firm producing the i^{th} variety of the domestic good. r is the distance for CBD, d_i the distance separating the i^{th} firm from the CBD and τ unit price commuting.

[2] And form the CBD.

[3] If they do so the price of land gets higher, so to attract them the firms must pay them higher wages.

In this context, if there are k domestic firms, the budget constraint of a worker living in r is

$$e_k Y_k - T_{k,i}(r) = R(r)\, s + z$$
$$\text{with } k \in \{g\ ,\ d\} \text{ and } i \in \{0,...,K\}$$

where $R(r)$ is the market price of one unit of housing in r.

The market price of housing in the city results from the aggregation of the bid-rent of all the workers living in the city. It is the maximum price that a worker is willing to pay for a unit of housing in a given location of the city at a given level of utility. As the city is open, utility levels are determined externally.

In the context of the arbitration described above, the bid-rent of a worker living in r is

$$\begin{aligned} \psi_k(r, u_k) &= \text{Max}_{z,s} \left\{ \frac{e_k Y_k - T_{k,j}(r) - z}{s} \mid U(z, s) = u_k \right\} \\ &= \text{Max}_s \left\{ \frac{e_k Y_k - T_{k,j}(r) - z(s, u_k)}{s} \right\} \end{aligned}$$
$$\text{with } k \in \{g\ ,\ d\} \text{ and } j \in \{0,...,K\}$$

First order conditions indicate that $\dfrac{e_k Y_k - T_{k,j}(r) - z(s, u_k)}{s}$ is the highest when s is solution of the following equation

$$-\frac{\partial z(s, u_k)}{\partial s} = \frac{e_k Y_k - T_{k,j}(r) - z(s, u_k)}{s}$$

By solving it we get $S_k(r, u_k)$, which is the optimal housing surface if the worker resides in r .

Thus, the worker residing in r bid-rent is :

$$\psi_e(r, u_e) = \frac{e_k Y_k - T_{k,j}(r) - z(S_k(r, u_k), u_e)}{S_k(r, u_k)}$$

A direct implication of the envelop theorem gives :

$$\frac{\partial \psi_k(r, u_k)}{\partial r} = \frac{-\frac{\partial T_{k,j}(r)}{\partial r}}{S_k(r, u_k)} < 0$$

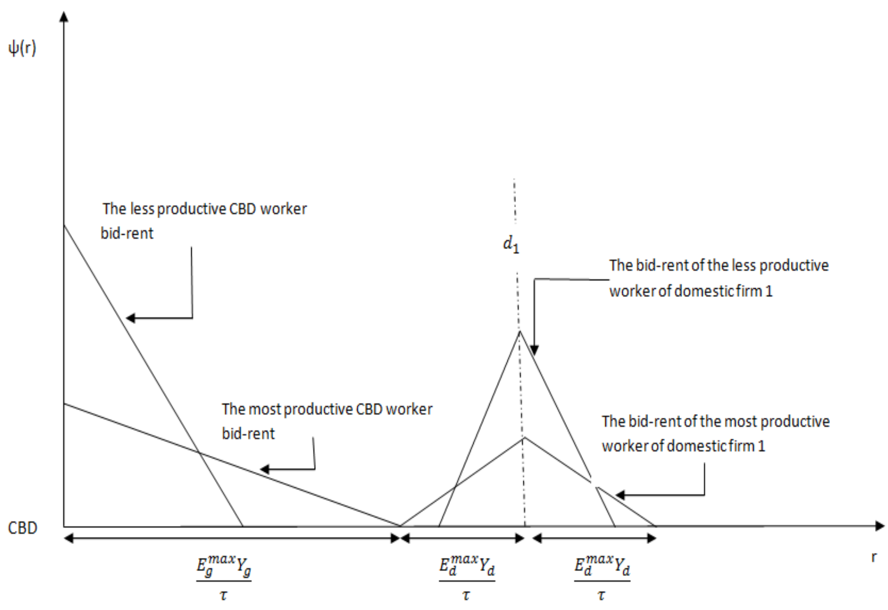
This relation states that the workers bid-rent decreases with the distance to their workplaces. Furthermore, as housing is a normal "service", $S_k(r, u_k)$ increases with income. Consequently, the bid-rent slope decreases with income. In terms of land use, this indicates that the lowest-paid workers prefer to pay a high price per square meter to get closer to their workcity even if they live in small dwellings, while the better-paid workers prefer to live far from their workcity and gain living space.

The decrease of the bid-rent with distance the workcity is reflected here, for the exporting firms workers, by the reduction of their bid-rent with distance to the CBD, and for the domestic firms workers, by the increase of theirs with r on the left of the firm that employs them and its diminution with r on its right (Cf Figure 1). Thus, the bid-rent of each firm workers intersect with those of the workers employed by the firm located on its right and those of the workers employed by the firm located on its left.

As a result, to ensure residential continuity while keeping workers' wages sufficiently low, in equilibrium, domestic firms locate so that the bid-rents of their highest-paid workers intersect at the location where they are equal to the agricultural rent R_a. For simplicity reasons, we set $R_a = 0$.

In this configuration, the intersections of the bid-rents with the x-axis determine the boundaries of the workers employed by each firm residence area. And since the remuneration of the exporting firms most productive worker is much higher than that of the domestic firms most productive one, the result in terms of bid-rent slopes indicates that residence area of the workers employed by the exporting firms is greater than that of domestic firms workers. The equilibrium city configuration is graphically depicted in Figure 1.

Figure 1: Land use equilibrium: heterogeneity across and within firms



2.4 Main testable prediction

The previous theoretical considerations lead us to make one prediction we then test.

Testable Prediction: Positive trade shock at the firm level increases wages and generates higher commuting distance at the worker level for exporters' employees.

3 Data

3.1 Data description

We use 4 datasets that cover the universe of French firms' employment in the Ile-de-France (IDF) region, surrounding Paris, from 2003 to 2008.

Worker wage and location data. We use is the *Déclaration Annuelle des Données Sociales* (DADS) files. This matched administrative employer-employee dataset comes from the mandatory reports by firms about their workforce each year and is made available to researchers by the INSEE (Institut National de la Statistique et des Etudes Economiques). While it covers the full universe of the private sector, we only use information on the individuals that worked once for a manufacturing firm. The unit of observation is a match between a firm (that can be identified) and an employee (that can not be identified) over two years. For each firm-worker observation, we have information about the individual gender, age, birth region, occupation (via 1 and 2-digit occupation code), annual gross and net earnings, the number of hours worked and the job status (full or part-time). Each worker-job observation also brings information about living and working places at the city level, using a 5-digit numerical code, similarly to a zip code. In our dataset, wages do not include potential monetary compensation for transport costs, ensuring that compensation for distance does not play a role in our exercise. We apply some restrictions on the dataset. First, we restrict the information to workers both living and working in the IDF region. Second, since workers cannot be followed for more than two consecutive years, we restrict our sample to workers that stay within the same firm across two consecutive years. We thus exclude job movers and restrict our sample to "job stayers". Our results shall thus be interpreted with care: our estimation measures the impact of trade shocks on the wage and location decision of a particular set of workers. Our results also neglect the possibility that firms may exit and close following trade shocks. We are however confident that producer exit is of low magnitude in our sample. As we can identify all firms in our sample, we find that on average, our sample contains more than 87% of all firms that we present once in the raw dataset. In our exercise, everything happens as if workers do not move across firms, and firms do not hire any additional workers.

Distance and time data. We merge this worker-level dataset with a unique dataset that provides information about the commuting distance and time for each origin-destination couple in the IDF. This dataset is provided by the *Société du Grand Paris*, and informs about the effective commuting distance and time between *zones* in the Ile-de-France region. We obtain a set of time-invariant bilateral commuting distances and times that covers all travels made within the region. We merge this information for each worker, given time-varying home and working cities.

Trade data. We also make use of firm-level trade data from the French customs to identify exporting firms and export shocks. This database reports the volume (in tons) and the value (in Euros) of exports for each CN8 product (European Union Combined Nomenclature at 8 digits) and destination, for each firm located on the French metropolitan territory. Some shipments are excluded from this data collection. Inside the EU, firms are required to report their shipments by product and destination country only if their annual export value exceeds the threshold of 150,000 Euros. For exports outside the EU, all flows are recorded unless their value is smaller than 1,000 Euros or one ton. Yet, these thresholds eliminate a very small share of the total exports. From this dataset, we only keep merchandise shipments, excluding agricultural and services exports. The raw dataset consists of 26,186,006 observations at the firm-year-destination-product level, that we aggregate into 7,110,894 observations at the firm-year-destination level and into 1,381,500 observations at the firm-year level. Once combined with the DADS in the Ile-de-France region, we obtain a dataset of 35 billion firm-worker-year observations over the 2003-2008 period, in which around 5,207,366 billions are employed in exporting firms (that represent 90% of the regional exports in the customs).

Balance-sheet data. We complete the picture using a balance-sheet dataset constructed from reports of French firms to the tax administration over the 2003-2008 period (Bénéfices Réels Normaux, BRN). This dataset contains information on the value added, total sales, capital stock, debt structure and other variables at the firm level. Importantly, this dataset is composed of all types of firms, including both small and large firms, since no threshold applies on the number of employees for reporting to the tax administration.

3.2 Final sample and descriptive statistics

Our exercice estimates the impact of trade activities on the worker-level decisions to change location. Our baseline sample only includes exporters' employees and contains 4,848,426 observations and covers 28,675 firms. Table 1 displays the descriptive statistics of our baseline sample.

Table 1: Descriptive statistics

	Obs.	Mean	Q1	Q3
Worker-level variables				
Mover Dummy	4,848,426	0.24	0	1
Distance (Private)	4,848,426	16.84	5.87	23.71
Time (Private)	4,848,426	25.10	11.790	35.37
Gross Year Wage	4,848,426	35616.56	18875	43769
Firm-level variables				
Exports	28,675	2080000	82971	3574283
Assets	28,675	262806	4123	43675
Employment	28,675	390.41	22	172
Apparent Labor Prod.	28,675	84.74	49	93.54

4 Empirical Strategy

4.1 General relationship

We identify the magnitude of the commuting distance elasticity to wage by estimating the following general equation:

$$\ln \text{Dist}_{ift} = \alpha_1 \ln \text{W}_{ift} + \alpha_2 C_{ft} + \alpha_3 C_{it} + \mathbf{FE} + \varepsilon_{ift} \quad (3)$$

where $Dist_{ift}$ is the commuting distance of any worker i in firm f in year t. The main dependant variable is the log hourly wage of any worker i in firm f - W_{ift} –. Parameter α_1 captures the elasticity of commuting distance to wage and is of prime interest to us. We include a set of firm-level controls, C_{ft} (log Assets, Log Apparent Labor Productivity, defined as total value added per worker), and a set of worker-level controls, C_{it} (i.e. individual observable worker characteristics: skilled/unskilled dummy, occupation group, age, gender.).

We include a set of fixed effects, **FE**, to account for unobserved heterogeneity across observations. We crucially include a worker fixed effect to (i) account for unobserved heterogeneity across workers and (ii) to properly identify our effect. When worker FE is included, parameter α_1 captures the effect of variation in wage on commuting distance coming from variations across years for a given worker. In other words, when we include this FE, we ensure that we estimate a within-worker effect instead of any across-workers effect. In that case, the estimated α_1 measures the elasticity of commuting distance to wage for a given worker, which is precisely what we want to measure. Then, this worker fixed effect absorbs all potential pre-existing trends the only remaining variation that we use stands across years thus absorbing As a counterpart, identification only arises from wage and commuting distance variation across years. Since we follow workers only for a two-year period, our estimate capture immediate change in location: our estimates are short-run results. We also include a year fixed effect, to account for unobserved heterogeneity across years, and both home-city FE and and work-city FE, to absorb any idiosyncratic effects. In some specification, we also add a firm × occupation (2-digit code), to control for any firm-specific trend in occupations. The estimation shall thus be understood as the effect within a firm-occupation layer across years for a worker.

4.2 Endogeneity concerns and IV

We could estimate equation (3) with an OLS estimator but there are many reasons to believe that the these estimates are biased. Even though the sample excludes job movers, we cannot exclude that wage and location decisions are simultaneously determined and that the OLS coefficient could be a biased estimator of the true parameter. To overcome endogeneity and to allow for causal interpretation of the coefficients, we use an instrumental variable (IV) strategy in a two-stage least square (2SLS) estimation. Our IV strategy has to identify sources of variations in wages that are truly exogenous to the firm and to workers, thus excluding the simultaneous impact on wages and distance. Identification of this relationship requires instruments that are orthogonal to firm-specific decisions. We thus make sure that our instruments are, by construction, exogenous to firm decisions and we will then test their validity.

First, wages are instrumented by the world import demand (WID hereafter) addressed to the products that are sold by the firm, so as to capture exogenous changes in trade conditions (Hummels et al., 2014; Berman et al., 2015; Mayer et al., 2016). This strategy thus focuses on demand shocks which, crucially, are orthogonal to firm-specific supply choices. Our baseline instrument is constructed using information about the foreign demand addressed to the firm using product and destination information. Specifically, we compute the sum of world imports in the products-destinations served by the firm in year t (using the BACI dataset) weighted by the share of each product-destination in the firm's total exports in year t (using the firm-level exports data). Weights are computed using the yearly share of the product-destination in the firms total exports. A product is defined at the 6-digit (HS6) level. More precisely, we define:

$$WID_{ft} = \sum_{s,j} \omega_{f,j,s,t} M_{jst} \tag{4}$$

where $\omega_{f,j,s,t}$ is the share of each product s and destination j in firm fs exports in year t, M_{jst} is the total value of imports for product s and destination j in year t, excluding France as potential exporter. By excluding French exports to this destination, we exclude sources of variations that originate in France and may be correlated with changes in the firm.

We include this instrument – WID – explicitly in the estimation. The following equation assesses the effect of exogenous changes in wages (occurring through variations in the instruments) on commuting distance, controlling for firm characteristics :

$$\ln \mathrm{W}_{ift} = \beta_0 + \beta_1 \ln WID_{ft} + \beta_2 C_{it} + \mathbf{FE} + \varepsilon_{ift} \tag{5}$$

$$\ln \mathrm{Dist}_{ift} = \alpha_1 \ln \widehat{\mathrm{W}_{ift}} + \alpha_2 C_{ft} + \alpha_3 C_{it} + \mathbf{FE} + \varepsilon_{ift} \tag{6}$$

where $\widehat{W_{ift}}$ is the predicted value of the log wage from equation (3). This strategy thus measures the effect of changes in wages, coming from trade shocks, on commuting patterns. We estimate these equations using a two-stage least-square estimation. In all estimations, standard errors are clustered at the 2-digit sector-year level.

5 Results

5.1 Main results

Our preferred estimation is thus obtained using an IV-2SLS estimation in which we instrument the (log) yearly wage by the (log) exogenous trade shocks.

Table 2 presents the result from the two stages of the estimation of equations (3) and (4). Each column corresponds to a different specification of equations (3) and (4. The upper panel of Table 2 presents the first-stage results of equation (3). We always estimate that changes in world demand are positively and significantly associated to changes in wages. In a second stage, we use the predicted value of wage from the first stage and estimate its impact of commuting distance. Since world import demand are exogenous to the firm and to workers, the estimates we obtain in the second stage are the causal impact of changes in wages on distance. In other words, we measure how much commuting distance is affected by variations in wages coming only from exogenous trade shocks.

Both panels in column (1) only include worker and year FE. Column (2) adds firm-year controls (log Assets, Log Apparent Labor Productivity) while column (3) adds home-city and work-city fixed effects to account for unobserved heterogeneity. Finally, in column (4) we include a set of 2-digit occupation FE. This specification is our preferred one and will be our benchmark estimation.

The top panel reports the first-stage results. Across specifications, coefficients all stand around 0.01 and precision is quite high. Since the specification is log-log, we estimate that a 10% increase in world import demand associates with a 0.4% increase in wage, ceteris paribus.

The bottom panel reports the second-step estimation results and we only report the coefficients on the log wage, which is of prime interest to us. Second step estimates confirm our main prediction. We always estimate a positive and significant coefficient to wage on commuting distance. Our preferred specification is in column (3) : we truly believe that including a large set of fixed effects is crucial to identify the effect. In this demanding estimation, results are consistent with a positive impact of variation in wages on commuting distance for a given worker across years. Our estimates mainly lie around 0.3. We estimate that a 10% increase in wage is associated at least to a 3% increase in commuting distance.

Quantitatively, our estimates are pretty small but can be explained by two reasons. First, we estimate contemporaneous changes in distance associated to changes in wage. For sure, there are many barriers leading to a slow and sticky change in commuting patterns, inducing a downward bias in the coefficient. Second, we only focus on the intensive margin of changes. The true estimate of the distance elasticity should include job stayers (our sample) and job movers (excluded in our sample). Our estimate is consequently negatively biased because these job movers may exhibit large changes in wages and in distance.

We then estimate how much this total effect on yearly wage can be decomposed in the number of hours worked (column (5)) and the hourly wage (column (6)). Second step estimates all confirm that the effect of world demand on distance is significant through change in the number of hours worked and not through hourly wage. First step estimates show that predicted changes in hours worked are significantly positive while hourly wages are not affected by trade exposition. Second step results confirm that exogenous trade shocks only affected commuting distance via the number of hours worked and not through hourly wage. This is consistent with extensive evidence regarding rigidities on the French labor market leading to small adjustments in wages following demand shocks.

5.2 Validity of instruments

Test for overidentifying restrictions. To check the validity of our strategy, we depart from our baseline estimations first by including a second instrument. we compute a firm-specific demand shock related to GDP. We thus compute a world demand –WD_{ft} hereafter – which is the weighted sum of GDP, weighted by the destination share in total exports of the firms: that capture total world demand arising from changes in GDP across destination countries:

$$WD_{ft} = \sum_j \omega_{fjt} GDP_{jt}. \tag{7}$$

Table 2: Baseline results: IV- 2SLS results

	(1)	(2)	(3)	(4)	(5)	(6)
			First stage			
Dependent Variable:	Year. Wage				Hours	Hr. Wage
WID	0.040^a	0.012^a	0.010^a	0.008^a	0.008^a	0.010
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)
R^2 (First Stage)	0.560	0.642	0.641	0.668	0.668	0.208
			Second stage Dependent Variable: Log Distance			
Log Wage	0.601^a	0.474^a	0.331^a	0.285^a	0.290^a	2.645
	(0.137)	(0.213)	(0.146)	(0.120)	(0.131)	(2.33)
Firm-Year Controls		x	x	x	x	x
Work city FE			x	x	x	x
House city FE			x	x	x	x
Occupation FE				x	x	x
Observations	4,848,426	4,848,426	4,848,426	4,848,426	4,848,426	4,848,426
R^2 (Second Stage)	0.650	0.671	0.675	0.680	0.731	0.532

Note: All columns include worker and year fixed effects. Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker.

While the world import demand used variations across products-destination, this measure uses variation solely across destinations and their GDPs. Our first stage estimation now becomes

$$\ln \mathrm{W}_{ift} = \beta_0 + \beta_1 \ln WID_{ft} + \beta_2 \ln WD_{ft} + \beta_3 C_{it} + \mathbf{FE} + \varepsilon_{ift} \tag{8}$$

while our second stage specification is unchanged.

Table 3 presents the estimation results and statistics to check the validity of our instruments. It replicates the structure of Table 2 in terms of dependent variables and specification, but it now includes two IVs. Including a second IV allows us to test for overidentifying restrictions. We check the validity of our instruments using two different tests. Robust to heteroskedasticity and clustering, Hansen J statistics for overidentifying restrictions are unable to reject our set of instruments. The F-stat form of the Kleibergen-Paap statistic as a test for weak instruments is also reported. All statistics are well above the critical values, confirming that our choice of instruments is statistically appropriate, on top of intuitive reasons explained above. Results using the IV-2SLS strategy are thus reliable since consequences of trade shocks are straightforwardly interpreted. Our shall be interpreted as causal effect since we made sure, by construction that variations in wage were exogenous to firms and workers' decisions.

Second-stage results are close to our baseline estimates. We always find a positive and significant relationship between yearly wage and commuting distance, of the same magnitude (around 0.3) as in the baseline estimates. Note that the precision of these estimates is higher, mainly because we included a second iv in the first stage. Once again, trade shocks only seem to affect the number of hours worked and not the wage rate, consistently with previous evidence.

Alternative weights in world demand. Second, we check the sensitivity of our results when using an alternative weighting scheme in our measure of world import demand. In particular, we compute the sum of world imports in the products-destinations served by the firm in year t weighted by the *average* share of each product-destination in the firm's total exports :

$$WID_{ft} = \sum_{s,j} \omega_{fjs} M_{jst} \tag{9}$$

where ω_{fjs} is the average share of each product s and destination j in firm fs exports over the total period, M_{jst} is the total value of imports for product s and destination j in year t, excluding France as potential exporter. We also compute a world import demand using product-destination *initial* shares

Table 3: Validity of the IV strategy

	(1)	(2)	(3)	(4)	(5)	(6)
	First stage					
Dependent Variable:	Year. Wage				Hours	Hr. Wage
WID	0.004^a	0.005^a	0.005^a	0.004^a	0.004^a	0.001^b
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
WD	0.032^a	0.008^a	0.001^a	0.004^a	0.004^a	0.002^a
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R^2 (First Stage)	0.548	0.549	0.448	0.547	0.549	0.234
	Second stage Dependent Variable: Log Distance					
Log Wage	0.486^a	0.394^a	0.306^a	0.227^a	0.190^a	0.076
	(0.071)	(0.071)	(0.072)	(0.059)	(0.061)	(0.064)
Firm-Year Controls		x	x	x	x	x
Work city FE			x	x	x	x
House city FE			x	x	x	x
Occupation FE				x	x	x
Observations	4,848,426	4,848,426	4,848,426	4,848,426	4,848,426	4,848,426
R^2 (Second Stage)	0.660	0.680	0.685	0.740	0.741	0.562
Hansen Stat.	3.409	3.306	3.265	1.856	1.954	2.435
Kleibergen-Paap stat.	27.646^b	27.236^b	27.097^b	$28{,}522^b$	23.661^b	7.248

Note: All columns include worker and year fixed effects. Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker.

Table 4: Alternative measure for world demand

	(1)	(2)	(3)	(4)	(5)	(6)
	First stage					
Dependent Variable:	Yearly Wage					
Weights	Average			Initial		
WID	0.036^a	0.016^a	0.017^a	0.016^a	0.008^a	0.011^a
	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)
R^2 (First Stage)	0.491	0.563	0.565	0.453	0.542	0.530
	Second stage					
	Dependent Variable: Log Distance					
Log Wage	0.385^a	0.298^a	0.296^a	0.253^a	0.190^a	0.134^a
	(0.081)	(0.083)	(0.091)	(0.041)	(0.042)	(0.045)
Firm-Year Controls	x	x	x	x	x	x
Work city FE		x	x		x	x
House city FE		x	x		x	x
Occupation FE			x			x
Observations	4,848,426	4,848,426	4,848,426	4,848,426	4,848,426	4,848,426
R^2 (Second Stage)	0.506	0.565	0.558	0.578	0.575	0.589

Note: All columns include worker and year fixed effects. Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker.

in exports as a weighting schemes:

$$WID_{ft} = \sum_{s,j} \omega_{fjst_0} M_{jst} \tag{10}$$

where t_0 corresponds the first year that the firms exports in our sample. Both measures check that time-variation of the foreign demand measure comes from the country-level imports by product, not from the firm-level weights.

Both stages' results using these two alternative measures of foreign demand are reported in table 4. All our baseline results are confirmed using alternative weighting schemes. All columns display positive, significant coefficients of log wage on commuting distance. Estimates are very close to our baseline results using the average weights, while coefficients are slightly lower in the case of initial shares.

5.3 Competing mechanisms

We check that our estimations do not capture alternative mechanisms. First, we check that our world import demand does not capture business cycles conditions that may also affect domestic demand. Then, we insulate our results from the import competition explanation.

Domestic sales. World business cycles conditions may affect both foreign and domestic sales. On top of that, at the firm level, foreign and domestic sales are well connected (Vannoorenberghe, 2012; Berman et al., 2015). So as to identify the proper impact of trade activities on wages and location choices, we insulate our results from domestic sales. The remaining effect thus identifies the change in distance solely coming from changes in revenues that remain orthogonal to domestic sales. We include domestic sales in our second stage estimation.

We measure domestic sales using two variables. First, balance-sheet data (BRN) reports the domestic sales values at the firm-year level. We include this dataset at the expense of many missing values in the sample. Second, to overcome this reduction in sample size, we compute a indirect domestic demand at the firm level, using an equivalent measure as our main instrument. We compute the sum of world imports coming from France for all products exported by the firm (using the BACI dataset) weighted by the share of each product in the firm's exports (from customs):

$$DomesticD_{ft} = \sum_{s,j} \omega_{fjst} M_{FR,st}. \tag{11}$$

This proxy thus provides a measure of demand, coming from France, for its products. It also allows us to recover a full set of domestic demands at the firm-year level, beyond what is reported in the balance-sheet data.

Table 5 provides the results of the IV regressions controlling for domestic demand. Columns 1 to 3 include the balance-sheet domestic sales measure while columns 4 to 6 use the proxy $DomesticD_{ft}$. Our baseline results are not affected by this additional control variable, independently of the measure. Magnitudes of the coefficients remain unchanged compared to the baseline coefficient, all around 0.3 in the most demanding specification. We also estimate a positive relationship between domestic demand and commuting distance: this is not surprising as both domestic and foreign demand shocks are likely to increase firm revenues.

Import competition. To be completed.

Table 5: Controlling for domestic sales

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Log Distance					
Domestic demand	Sales			Predicted demand		
Log Wage	0.501^a	0.274^a	0.231^a	0.385^a	0.290^a	0.251^a
	(0.137)	(0.213)	(0.146)	(0.120)	(0.131)	(0.149)
Log Domestic demand	0.010^a	0.011^a	0.007^a	0.005^a	0.008^a	0.006^a
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Firm-Year Controls	x	x	x	x	x	x
Work city FE		x	x		x	x
House city FE		x	x		x	x
Occupation FE			x			x
Observations	2,798,092	2,798,092	2,798,092	4,848,426	4,848,426	4,848,426
R^2 (Second Stage)	0.579	0.737	0.737	0.746	0.747	0.782

Note: All columns include worker and year fixed effects. Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker. Domestic demand measures are defined in the main text.

5.4 Locations choices

We now replicate the same type of estimation but focusing an alternative dependent variables, all related to the location decisions of workers. In particular, since our strategy relies on a diff-in-diff estimation, our data allows us to study the location choice of the workers. By including information about city characteristics in our dataset, we can interpret the choice of workers in terms of amenities and economic characteristics of new home cities compared to former home cities.

Table 6 provides our second stage estimation of the following equation:

$$CityCaract_{ct} = \alpha_1 \ln \widehat{\mathrm{W}}_{ift} + \alpha_2 C_{ft} + \alpha_3 C_{it} + \mathbf{FE} + \varepsilon_{ift} \tag{12}$$

where $CityCaract_{ct}$ represents time-invariant city characteristics in terms of employment, density, revenues... By comparing city characteristics, that do not change over time, across years for a given worker, coefficient α captures whether change in home cities correlates with differences in city characteristics. Since city characteristics are time-invariant, we drop the cities FE from our benchmark fixed effect structure. First stage regression remains unchanged (see equation (**??**)).

We consider many city characteristics in our exercise. Column (1) focuses on the impact of changes in wages on the distance to the center of Paris. We compute this distance as the distance with respect to the 75001 code, that corresponds to the first arrondissement of Paris. We thus obtain a set of unilateral distances at the city level. Previous results supported that exogenous changes in wages led

Table 6: Location choice

	(1)	(2)	(3)	(4)	(5)
Dependent Var.	Distance to Paris	Revenue 2002	Workers Density	Firms Density	Exporters Density
Log Wage	-0.039^a	0.128^a	0.019^a	0.018^a	0.035^a
	(0.008)	(0.011)	(0.004)	(0.005)	(0.005)
Firm-year controls	x	x	x	x	x
Worker FE	x	x	x	x	x
Occupation FE	x	x	x	x	x
Observations	4,848,426	4,848,426	4,848,426	4,848,426	4,848,426
R^2	0.650	0.612	0.648	0.432	0.539

Note: All columns include worker and year fixed effects. Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker.

workers to move away from their firms. The negative and significant coefficient in column (1) suggests, that that, on top of this, workers moved closer to the center of Paris. The coefficient here comes from the difference between distance to Paris of home city in t and in $t-1$ for any worker.

Column (2) estimates the correlation between changes in wages and the average revenue of cities in 2002, which is out of our sample. We use the full DADS dataset (see the Data section) and compute the average yearly revenue for each (home) city. We estimate that on average, workers that experienced increases in wages tend to move to richer cities, compared to their former home city.

Does density matter for the location choice? We measure worker, firm and exporting density as the total number of workers, firms (both from the DADS dataset) and exporting firms (from the Customs dataset) per square kilometre in each city. City size data comes from the CORINE Land Cover (CLC). Columns (3) to (5) state that workers that experienced increases in wages tend to move to cities with higher workers' density (column (3)), firms density (column (4)) and exporting firms density (column (5)), compared to their former home city. We thus infer from this set of results that the existence of trade activities tends to make workers move away from their firm, towards rich cities, closer to the region center and denser cities.

6 Conclusions

In the present paper, we have studied the impact of trade-related wage variations on location decision of exporting firms workers. We have provided a theoretical model capturing the spatial sorting of workers following changes in wages. The model emphasizes a direct effect of trade shocks on exporters' employees – through a wage effect –, that we directly test empirically. We provided micro-level support for this set of results. Using IV-2SLS strategies, we estimate a causal positive worker-level commuting distance elasticity to wage, using datasets covering the universe of Ile-de-France (surrounding Paris). We identified trade shocks that are exogenous to firms and workers by computing, using disaggregated customs data, the world import demand that is addressed to the firm.

While this paper provides a clear identification of the wage-commuting relationship, it also opens research lines that can be investigated using the same framework. In particular, our results have important implications for aggregate impacts of trade at the city level on its organisation, its structures and agglomeration forces. A proper identification however requires to assess the impact of trade on domestic workers. We leave this investigation for future research.

References

Alonso, William, *Location and Land Use*, Cambridge: Harvard University Press, 1964.

Berman, Nicolas, Antoine Berthou, and Jérôme Héricourt, "Export dynamics and sales at home," *Journal of International Economics*, 2015, *96* (2), 298–310.

Bombardini, Matilde, Gianluca Orefice, and Maria D. Tito, "Does Exporting Improve Matching? Evidence from French Employer-Employee Data," NBER Working Papers 21225, National Bureau of Economic Research, Inc May 2015.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

Davidson, C. and Nicholas Sly, " Trade and the labor market: Recent developments and new frontiers," 2012.

Davidson, Carl, Fredrik Heyman, Steven Matusz, Fredrik Sjoholm, and Susan Chun Zhu, "Liberalized Trade and Worker-Firm Matching," *American Economic Review*, May 2012, *102* (3), 429–434.

Fujita, Masahisa, Jacques-Francois Thisse, and Yves Zenou, "On the Endogeneous Formation of Secondary Employment Centers in a City," *Journal of Urban Economics*, May 1997, *41* (3), 337–357.

Grossman, Gene, "Heterogeneous workers and international trade," *Review of World Economics (Weltwirtschaftliches Archiv)*, June 2013, *149* (2), 211–245.

Hummels, David, Rasmus Jï£¡rgensen, Jakob Munch, and Chong Xiang, "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data," *American Economic Review*, June 2014, *104* (6), 1597–1629.

i Puigarnau, Eva Gutierrez, Ismir Mulalic, and Jos N. van Ommeren, "Do rich households live farther away from their workplaces?," *Journal of Economic Geography*, 2016, *16* (1), 177–201.

Mayer, Thierry, Marc J. Melitz, and Gianmarco I.P. Ottaviano, "Product Mix and Firm Productivity Responses to Trade Competition," Working Paper 22433, National Bureau of Economic Research July 2016.

Mulalic, Ismir, Jos N. Van Ommeren, and Ninette Pilegaard, "Wages and Commuting: Quasi-natural Experiments' Evidence from Firms that Relocate," *Economic Journal*, 09 2014, *124* (579), 1086–1105.

Muth, Richard F, *Cities and Housing*, Chicago: University of Chicago Press., 1969.

Straszheim, Mahlon, "The theory of urban residential location," in E. S. Mills, ed., *Handbook of Regional and Urban Economics*, Vol. 2 of *Handbook of Regional and Urban Economics*, Elsevier, 1987, chapter 18, pp. 717–757.

Timothy, Darren and William C. Wheaton, "Intra-Urban Wage Variation, Employment Location, and Commuting Times," *Journal of Urban Economics*, September 2001, *50* (2), 338–366.

Vannoorenberghe, G., "Firm-level volatility and exports," *Journal of International Economics*, 2012.

Zenou, Yves, *Urban Labor Economics* number 9780521698221. In 'Cambridge Books.', Cambridge University Press, March 2009.