

Pokrovsky, Dmitry; Behrens, Kristian

**Conference Paper**

## Unequal cities: Self-selection, matching, and the distribution of income

54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

Suggested Citation: Pokrovsky, Dmitry; Behrens, Kristian (2014) : Unequal cities: Self-selection, matching, and the distribution of income, 54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

This Version is available at:

<http://hdl.handle.net/10419/124410>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Unequal cities: Self-selection, matching, and the distribution of income \*

Kristian Behrens<sup>†</sup>    Dmitry Pokrovsky<sup>‡</sup>

October 31, 2013

## Abstract

We develop a model of a city populated by heterogeneous agents. Agents self-select into entrepreneurship, and entrepreneurs set up firms which hire workers. We characterize the equilibrium matching between firms and workers, as well as the within-city assignment of agents to locations. We then explore the implications of city size and the characteristics of the underlying skill distribution for selection into entrepreneurship, rent gradients, and city-wide inequality in disposable incomes. We also derive several testable predictions and confront them with the data.

**Keywords:** cities; income inequality; firm-worker matching; self-selection; rent gradients

**JEL Classification:** R10; R12; R13

## 1 Introduction

The sorting of heterogeneous workers across cities is empirically important (Combes, Duranton, Gobillon, and Roux, 2012) and has recently been reconsidered from a theoretical perspective (Behrens, Duranton, and Robert-Nicoud, 2013; Davis and Dingle, 2013). Sorting of workers, as well as occupational selection, is known to increase average productivity, to be associated with city size, and to eventually lead to a more skewed distribution of incomes across agents in large urban areas (Behrens and Robert-Nicoud, 2013). Until now, there are no models that jointly look at agglomeration, sorting, and ‘non-trivial’ selection

---

\*Financial support from the CRC Program of the Social Sciences and Humanities Research Council (SSHRC) of Canada – for the funding of the *Canada Research Chair in Regional Impacts of Globalization* – and from the *Center for Market Studies and Spatial Economics*, NRU-HSE, are gratefully acknowledged.

<sup>†</sup>Canada Research Chair in Regional Impacts of Globalization, Département des Sciences Économiques, Université du Québec à Montréal (UQAM), Canada; CIRPÉE; and CEPR. E-mail: [behrens.kristian@uqam.ca](mailto:behrens.kristian@uqam.ca)

<sup>‡</sup>Center for Market Studies and Spatial Economics, National Research University Higher School of Economics, Russia. E-mail: [dm.pokrovsky@gmail.com](mailto:dm.pokrovsky@gmail.com)

(selection in Behrens, Duranton, and Robert-Nicoud, 2013, is trivial and does not depend on city size, whereas in Behrens and Robert-Nicoud, 2013, there is no sorting). Furthermore, whereas the ‘macro’ aspects of sorting across cities has been the focus on the literature, much less is known about sorting of heterogeneous agents within cities. Although this aspect has been considered in the theory of local public goods and in local public finance, the city-wide implications have not been derived. For example, there are – to the best of our knowledge – no results in the literature that link the distribution of income and income inequality to land rent profiles. Yet, the latter are essential if one wants to assess ‘real’ income inequality in cities, i.e., controlling for the locational profile of agents and the fact that there is a systematic link between incomes and land rent. Taking into account these various aspects of within-city sorting and selection, the aim of this project is to investigate in detail the locational equilibrium in cities and to link various aspects of the income (or skill) distribution to city-wide outcomes like the land-rent gradient or ‘real’ income inequality.

We develop a model of a city populated by heterogeneous agents. Agents self-select into entrepreneurship, and entrepreneurs set up firms which hire workers. We characterize the equilibrium matching between firms and workers, as well as the within-city assignment of agents to locations in the spirit of Beckmann (1969). We then explore the implications of city size and the characteristics of the underlying skill distribution for selection into entrepreneurship, rent gradients, and city-wide inequality in disposable incomes. Our contribution, as compared to others, is that we explain the city structure via a self-selection mechanism of agents into entrepreneurship and workers and that we look at the formation of the bid-rent for agents who differ with respect to income.

Agents self-select into occupations based on their earnings, as in Lucas (1978). More precisely, we assume that highly productive agents have a comparative advantage in entrepreneurship, so that there exists some endogenously determined unique cutoff  $\hat{\varphi}$  that separates workers from entrepreneurs:  $\Omega_w = [\underline{\varphi}; \hat{\varphi}]$  (workers) and  $\Omega_e = [\hat{\varphi}; \bar{\varphi}]$  (entrepreneurs). Workers are hired by entrepreneurs and are paid a match-specific wage,  $w(\varphi_e, \varphi_w)$ , whereas entrepreneurs are the residual claimants to their firms’ profits,  $\pi(\varphi_e, \varphi_w)$ . All agents live in a linear city that stretches out on the interval  $[0; l]$ , where  $l$  denotes the endogenously determined city fringe. The central business district (CBD), where all firms are concentrated, is located at 0. Both workers and entrepreneurs commute to the CBD for work. Commuting entails costs, which we model parsimoniously using an ‘iceberg’ specification: if an agent in location  $l$  has income  $I(\varphi)$ , his commuting costs are given by  $I(\varphi)(1 - \exp(-\tau l))$ . Thus, the income net of commuting costs is  $I(\varphi)\exp(-\tau l)$ . The parameter  $\tau > 0$  captures ‘urban frictions’ that arise in the city (see, e.g., Lucas and Rossi-Hansberg, 2002; Behrens, Mion, Murata, and Südekum, 2012).

In our model, agents differ by income  $I(\varphi)$ . We make no assumptions on the income distribution, except that  $I'(\varphi) > 0$ : more productive agents earn higher incomes. Since commuting costs are paid as a fraction of income, it follows that richer agents – equivalently, agents with a higher skill level – will want to locate closer to the CBD to minimize their costs. We thus have to find the

spatial distribution of agents in the city (Beckmann, 1969). Formally, we solve a locational assignment problem, i.e., we have to find the mapping of agents,  $\varphi$ , to city locations,  $l$ , such that every agent picks his preferred location, i.e., no agent has any incentives to relocate. We will first look at the simple case with a fixed lot size assumption, in which case maximizing utility is equivalent to maximizing disposable income – income net of land rent and commuting costs. We also will look at the more complicated case where the consumption of land is endogenous.

We anticipate to obtain the following results. First, we will derive the comparative statics of the equilibrium variables with respect to exogenous parameters of the model like total population, commuting costs, and various moments of the talent distribution of the population. We expect that reducing commuting costs will lead to an increase in the density in a city, and that it will increase the share of entrepreneurs in the city. A larger share of entrepreneurs leads to tougher selection, which should magnify income inequality in the city. We have no a priori intuition for the direction of change in ‘real’ income inequality. Furthermore, a distribution of talent that is more skewed towards highly talented workers should increase the steepness of the rent gradient towards the city center. The reason is that more productive agents live closer to the center and compete for land there, and that a larger mass of highly talented agents will increase competition for land towards the center. Depending on how fast land prices go up for the rich compared to land prices for the poor, ‘real’ income inequality may a priori rise or fall as the distribution of talent gets more unequal within the city. Ideally, we would also integrate consideration on land-use restrictions into our model, but this extension may be intractable in our complex framework, in which case we will keep it for future research.

To achieve the goals set out in this project, we will use the following approaches: (i) Self-selection mechanisms from occupational choice models; (ii) construction of the bid-rent function from city models in the wake of Beckmann (1969); and (iii) standard microeconomic methods linked to the investigation of heterogeneous firms’ models.

We will start by focusing on a closed city, and we will eventually try to extend the model to a multi-city setting.

This project is largely theoretical, but we expect that not all results will be derivable using pencil and paper only. Hence, we expect to use numerical methods to simulate the model, and to establish some of the comparative static results with respect to land rents and to ‘real’ income inequality. The numerical methods we will use are expected to be very standard and they can be implemented using standard software such as Mathematica, Matlab, or some C++ customized code.

While we have no clear plan yet, we ideally also want to empirically test some of the propositions derived from the model. In particular, we plan to use US data on income inequality and rent gradients to test whether a more skewed income distribution drives up rents towards the city center, and whether this reduces ‘real’ income inequality. We also plan to eventually make use of French data that is currently assembled by Combes, Duranton and Gobillon, (2012) and

that would allow us to test some of the within-city predictions of the model.

## 2 General setup

Consider an economy with  $L$  agents. Agents differ in their productivity,  $\phi$ , and endogenously choose between setting up a firm ('entrepreneurs') or being hired by firms ('workers'). Each entrepreneur produces a distinct variety of a horizontally differentiated consumption good. The productivity,  $\phi$ , has a continuously differentiable cumulative distribution function  $F(\cdot)$  on  $\Omega = [\underline{\phi}, \bar{\phi}]$ . An agent with productivity  $\phi$  produces  $1/(\phi\xi)$  units of a differentiated good, where  $\xi$  is an endogenously determined characteristic that depends on his match with a firm – if he is a worker – or his match with workers – if he is an entrepreneur. We subscript variables related to workers with  $w$ , and variables related to entrepreneurs with  $e$ . Because of matching between workers and entrepreneurs,  $\phi_e = \phi_e(\phi_w)$  and  $\phi_w = \phi_w(\phi_e)$ . We will make precise the matching procedure later.

Agents self-select into occupations based on their earnings, as in Lucas (1978). More precisely, we assume that highly productive agents have a comparative advantage in entrepreneurship, so that there exists some endogenously determined unique cutoff  $\hat{\phi}$  that separates workers from entrepreneurs:

$$\Omega_w = [\underline{\phi}; \hat{\phi}] \text{ (workers)} \quad \text{and} \quad \Omega_e = [\hat{\phi}; \bar{\phi}] \text{ (entrepreneurs)}. \quad (1)$$

Workers are hired by entrepreneurs and are paid a match-specific wage,  $w(\phi_e, \phi_w)$ , whereas entrepreneurs are the residual claimants to their firms' profits,  $\pi(\phi_e, \phi_w)$ .

### 2.1 Preferences and urban structure

All agents live in a linear city that stretches out on the interval  $[0; \bar{l}]$ , where  $\bar{l}$  denotes the endogenously determined city fringe.<sup>1</sup> The central business district (CBD), where all firms are concentrated, is located at 0. Both workers and entrepreneurs commute to the CBD for work. Commuting entails costs, which we model parsimoniously using an 'iceberg' specification: if an agent in location  $l$  has income  $I(\phi)$ , his commuting costs are given by  $I(\phi)(1 - e^{-\tau l})$ . Thus, the income net of commuting costs is  $I(\phi)e^{-\tau l}$ . The parameter  $\tau > 0$  captures 'urban frictions' that arise in the city (see, e.g., Lucas and Rossi-Hansberg, 2002; Murata and Thisse, 2005; Behrens, Mion, Murata, and Südekum, 2012).

Each agent inelastically consumes one unit of housing, which is distributed with a cumulative distribution function  $G(\cdot)$  in the city. We assume that this distribution is exogenously given. In that case,  $L \equiv \int_0^{\bar{l}} dG(\ell)$  because of unit lot size. Since agents differ by productivity  $\phi$ , it will be useful to index agents by that productivity parameter. The consumer problem of a type- $\phi$  agent, when he lives at location  $l$ , is given by

<sup>1</sup>We could consider a city that is symmetric on  $[-\bar{l}, \bar{l}]$ . Our results would be unchanged.

$$\max_{x(\cdot), \nu \in \Omega_e} \int_{\Omega_e} u(x(\nu)) dF(\nu) \quad \text{s.t.} \quad \int_{\Omega_e} p(\nu)x(\nu)dF(\nu) + R(l) = I(\phi)e^{-\tau l}, \quad (2)$$

where  $I(\phi)$  denotes the consumer's gross income, which is given by the wage if she is a worker and by the profit if she is an entrepreneur; where  $R(l)$  denotes the land rent in location  $l$ ; where  $p(\nu)$  and  $x(\nu)$  denote the price and the consumption of a variety produced by a type- $\nu$  entrepreneur, respectively; and where  $e^{-\tau l}$  is the share of income remaining after paying commuting costs.

The first-order conditions for utility maximization are given by:

$$\frac{u'(x(\nu))}{\lambda(\phi)} = p(\nu), \quad \forall \nu \in \Omega_e \quad (3)$$

where we have made explicit the fact that the marginal utility of income – the Lagrange multiplier  $\lambda(\phi)$  – is consumer specific. Borrowing the notation of Zhelobodko, Kokovin, Parenti, and Thisse (2012), we define the elasticity of inverse demand,  $r$ , as follows:

$$r(\nu) = -\frac{u''(x(\nu))x(\nu)}{u'(x(\nu))}. \quad (4)$$

It can be verified that  $r$  is also equal to the Lerner index. Using (3), the aggregate quantity sold by a type- $\nu$  firm is given by:

$$X(\nu) = L \int_{\Omega} (u')^{-1}(\lambda(\phi)p(\nu)) dF(\phi), \quad (5)$$

which generally depends on the distribution of incomes across consumers via the distribution of Lagrange multipliers.<sup>2</sup> We can write the aggregate demand as a function of a ‘generalized multiplier’ – an aggregate market statistic – when the inverse of the marginal utility is either additively or multiplicatively quasi-separable (Behrens and Murata, 2007; Zhelobodko *et al.*, 2012). In that case, we have

$$X(\nu) = Lf(p(\nu), \Lambda), \quad \text{where} \quad \Lambda \equiv \int_{\Omega} g(\lambda(\phi)) dF(\phi) \quad (6)$$

and where  $f$  and  $g$  are some functions. One big advantage of utilities that allow for generalized multipliers is that we only need to solve for the aggregate market statistic, not for the whole distribution of Lagrange multipliers. We will see later specific examples of utility functions that satisfy these properties. For now, we consider the general case.

---

<sup>2</sup>The market need not be fully covered in our specification. Firms may find it profitable to price low-income consumers out of the market if the elasticity of demand for the remaining consumers falls sufficiently quickly.

## 2.2 Locational assignment and land rent

In our model, agents differ by income  $I(\phi)$ . For now, we make no assumptions on the income distribution, except that  $I'(\phi) > 0$ : more productive agents earn higher incomes. Since commuting costs are paid as a fraction of income, it follows that richer agents – equivalently, agents with a higher  $\phi$  – will want to locate closer to the CBD to minimize their costs. We thus have to find the spatial distribution of agents in the city (Beckmann, 1969). Formally, we solve a *locational assignment problem*, i.e., we have to find the mapping of agents,  $\phi$ , to city locations,  $l$ , such that every agent picks his preferred location, i.e., no agent has any incentives to relocate. Under the fixed lot size assumption, maximizing utility is equivalent to maximizing *disposable income*, i.e., income net of land rent and commuting costs. Formally, the locational assignment is such that:

$$\mu : \phi \in \Omega \rightarrow l \in [0, \bar{l}], \quad l \in \operatorname{argmax}_l \text{DI} \equiv I(\phi) e^{-\tau l} - R(l). \quad (7)$$

Observe that since  $(\partial^2 \text{DI})/(\partial l \partial I) = -\tau e^{-\tau l} < 0$ , the disposable income of richer agents increases faster as they move towards the center than that of poorer agents. Since land is allocated to the highest bidder, this implies that richer agents will live closer to the CBD as they can outbid the other agents: there is positive assortative matching (PAM) between incomes and locations close to the center. Put differently, the bid rent function of higher income types is steeper than that of lower income types. Because of unit lot sizes, the locational assignment is then simple to derive: the  $x\%$  of the locations closest to the CBD will be occupied by the  $x\%$  of the richest agents. As income increases with productivity  $\phi$ , we thus have the assignment

$$\int_{\underline{\phi}}^{\varphi} dF(\phi) = \int_{\mu(\varphi)}^{\mu(\underline{\phi})} dG(\ell), \quad (8)$$

for any  $\varphi \in [\underline{\phi}, \bar{\phi}]$  and for any cumulative distribution of land  $G(\cdot)$  in the city.

The rent schedule is derived from the optimal location of a type- $\phi$  agent with income  $I(\phi)$  using the following conditions:

$$\frac{d}{dl} [I(\phi) e^{-\tau l} - R(l)] = 0 \quad \text{and} \quad r(\bar{l}) = 0,$$

where the latter condition is a normalization of land rent to zero at the city fringe. The *conditional* rent gradient – which depends on  $\phi$  – is given by

$$R'(l | \phi) = -\tau e^{-\tau l} I(\phi). \quad (9)$$

The *unconditional* rent gradient makes use of the assignment of types to locations. Hence, the profile of the rent schedule depends on the spatial distribution of incomes in the city:

$$R(l) = \tau \int_0^{\bar{l}} e^{-\tau \ell} I(\mu^{-1}(\ell)) d\ell - \tau \int_0^l e^{-\tau \ell} I(\mu^{-1}(\ell)) d\ell, \quad (10)$$

where we have made use of  $R(\bar{l}) = 0$  and the assignment  $\mu$  of types to locations.

Two remarks are in order. First, contrary to urban models without heterogeneity, the utility levels of agents are *not a priori equalized* across locations. This is clear from the fact that disposable income varies across space, which is then enough to generate differences in utility since only disposable income matters for the consumption level of the differentiated good. Second, without imposing further conditions on the distribution of incomes, we cannot obtain general comparative static results as to how changes in the distribution of incomes affect the rent gradient. We will see later – using specific parametrizations – that more unequal cities generally have a steeper rent gradient close to the center than more egalitarian cities. The reason for this is that in cities with a larger mass of rich consumers, competition for land is fiercer closer to the center. Since richer agents pay much higher land rents in richer cities, this implies that inequality in disposable incomes is lower in such cities (Moretti, 2012).

### 2.3 Production and wage premium

Entrepreneurs maximize their operational profit, i.e., the difference between revenue and variable cost. Each entrepreneur must hire workers, and the *productivity of the firm is match specific*, i.e., it depends on the type of worker hired.<sup>3</sup> When an entrepreneur with productivity  $\phi_e$  hires workers with productivity  $\phi_w$ , the firm’s productivity is given by  $\phi_e \times \phi_w$ .

Contrary to heterogeneous firms models à la Melitz (2003), we assume that production entails no fixed costs. Since entrepreneurs are the residual claimants to firms’ profits, their income is equal to the operational profit that they earn. That operational profit – conditional on hiring workers of type  $\phi_w$  and paying them wages  $w(\phi_e, \phi_w)$  – is given by:<sup>4</sup>

$$\pi(\phi_e, \phi_w) = \left[ p(\phi_e) - \frac{w(\phi_e, \phi_w)}{\phi_e \phi_w} \right] X(\phi_e), \quad (11)$$

where  $X(\phi_e)$  is defined in (5). Let  $c(\phi_e, \phi_w) \equiv \frac{w(\phi_e, \phi_w)}{\phi_e \phi_w}$  denote the firm’s variable cost. From the first-order conditions for profit maximization, we get the profit-maximizing price

$$p(\phi_e, \phi_w) = \frac{c(\phi_e, \phi_w)}{1 - r(\phi_e, \phi_w)},$$

where  $r$  is given by equation (4). Let  $\pi^*(\phi_e, \phi_w)$  denote the maximized operating profits. Firms hire workers optimally, i.e., they select workers with productivity

---

<sup>3</sup>Behrens, Duranton, and Robert-Nicoud (2012) also consider productivities that depend multiplicatively on two terms: skill and luck. However, in their model there is no matching between heterogeneous agents. Skills are innate, and luck is a random shock that is revealed once location choices are made.

<sup>4</sup>We assume that firms hire only a single type of worker. Under positive assortative matching with a continuum of firms and a continuum of workers, we can always consider that a one-to-one assignment holds.

$\phi_w$  such that

$$\frac{d\pi^*(\phi_e, \phi_w)}{d\phi_w} = 0.$$

Workers select the firms that pay them the highest wages, which implies that the equilibrium wage schedule  $w(\phi_e, \phi_w)$  must support the assignment of workers to firms (Sattinger, 1993, discusses in detail the differential assignment problem). Dropping arguments from functions to alleviate notation, the matching between firms and workers then solves the following differential equation:

$$\frac{d\pi^*(\phi_e, \phi_w)}{d\phi_w} = \frac{d\pi^*}{dp} \left( \frac{dp}{dc} \frac{dc}{d\phi_w} + \frac{dp}{dr} \frac{dr}{d\phi_w} \right) + \frac{d\pi^*}{dc} \frac{dc}{d\phi_w} = \frac{d\pi^*}{dc} \frac{dc}{d\phi_w} = 0,$$

where the second equality comes from firms' profit maximization ( $d\pi^*/dp = 0$ ). It can readily be verified that we have

$$\frac{dc}{d\phi_w} = \frac{w}{\phi_w^2 \phi_e} \left( \frac{dw}{d\phi_w} \frac{\phi_w}{w} - 1 \right). \quad (12)$$

As can be seen from (12),  $\frac{dw}{d\phi_w} \frac{\phi_w}{w} = 1$  determines the unique wage schedule that sustains the equilibrium assignment of workers to firms. This implies that the wage schedule is of the form  $w(\phi_w) = A\phi_w$ , where  $A$  is some positive constant. In what follows, we impose the normalization condition  $w(\phi) = 1$ , i.e., the lowest-skilled workers' wages are taken as the numéraire. Under that condition, we can pin down the wage schedule that firms have to pay to workers of the optimal type they want to hire:

$$w(\phi_w) = \frac{\phi_w}{\phi}. \quad (13)$$

A few comments are in order. First, as can be seen from equation (13), wages are linearly increasing in workers' productivity. Consequently, more productive workers are paid the full value of their additional productivity by the firm. This result is independent of consumers' preferences in our model, and it is driven by the competitive and frictionless assignment of workers to firms.<sup>5</sup> Second,  $w(\phi_w)$  can be directly interpreted as the worker's skill premium when he is of type  $\phi_w$ . There is no skill premium for the lowest-skilled workers, and then the skill premium linearly rises with workers' productivity. Third, the skill premium does not directly depend on city size or the distribution of productivities. Larger cities do not pay higher wages to workers. This is because there are no agglomeration economies in our model. Having agglomeration economies like input sharing, or better matching in the urban labor market, would allow to have wages that are increasing in city size. Sorting along skills as in Behrens *et al.* (2012) would also allow for higher wages in larger cities by changing the skill mix of workers. Note, however, that in our model larger cities still provide different

<sup>5</sup>If wages were set differently, e.g., via bargaining or some other mechanism, the pass-through of workers' productivity to wages would be less than one.

incentives for self-selection into entrepreneurship. Consequently, the selection cutoff and the distribution of income between workers and entrepreneurs will usually differ depending on city size. Last, although workers' nominal incomes do not depend on city size, their disposable income will depend on city size. Larger cities have higher land rents, which implies that workers' disposable income will be lower in larger cities. Since larger cities also carry a larger array of consumption goods, the net impact on welfare is a priori unclear.

## 2.4 Selection and assignment of workers to firms

Agents self-select into occupations based on the returns they can earn. As shown before, wages are strictly increasing in workers' productivity,  $\phi_w$ . We assume that more productive agents have a comparative advantage in entrepreneurship, i.e., profits are strictly increasing in entrepreneurial productivity,  $\phi_e$ , and they increase at a faster rate than wages:

$$(\pi^*)'(\phi) \geq w'(\phi), \quad \forall \phi \in [\underline{\phi}, \bar{\phi}].$$

Provided that  $\pi(\underline{\phi}) < w(\underline{\phi}) = 1$  and  $\pi(\bar{\phi}) > w(\bar{\phi}) = \bar{\phi}/\underline{\phi}$ , there then exists by continuity a unique cutoff productivity level,  $\hat{\phi}$ , that satisfies:

$$\pi^*(\hat{\phi}) = w(\hat{\phi}).$$

In words, profits increase at a faster rate than wages and they cross the wage schedule a single time, which makes sure that highly productive agents are entrepreneurs whereas less productive agents are workers. This also implies – via the locational assignment – that entrepreneurs are located closer to the center, whereas workers live further away.

We assume that maximized profits – taking wages as given by firms – are supermodular in entrepreneurs' and workers' productivity:

$$\begin{aligned} \frac{\partial^2 \pi^*}{\partial \phi_e \partial \phi_w} &= \frac{X(\phi_e)}{\phi_e \underline{\phi} (1-r)^2} \left[ -\frac{1}{\phi_e} \frac{\partial r}{\partial \phi_w} + \frac{\partial^2 r}{\partial \phi_w \partial \phi_e} + \frac{2}{1-r} \frac{\partial r}{\partial \phi_w} \frac{\partial r}{\partial \phi_e} \right] \\ &+ \frac{\partial p}{\partial \phi_e} \frac{\partial r}{\partial \phi_w} \frac{1}{(1-r)^2} L \int_{\Omega} ((u')^{-1})'(\lambda(\phi)p(\phi_e)) \lambda(\phi) dF(\phi) > (15) \end{aligned} \quad (14)$$

A sufficient condition for (15) to hold, so that profits are supermodular in both productivities, is that

$$\frac{\partial r}{\partial \phi_e} \leq 0, \quad \frac{\partial r}{\partial \phi_w} \leq 0, \quad \text{and} \quad \frac{\partial^2 r}{\partial \phi_w \partial \phi_e} \geq 0, \quad (16)$$

with at least one inequality being strict. It can readily be verified that under those conditions,  $\frac{\partial p}{\partial \phi_e} < 0$ , i.e., more productive entrepreneurs charge lower prices. Observe that the conditions (16) are not overly restrictive. In particular, the condition on the cross-derivative is always satisfied in our setup, because cost  $c(\phi_e, \phi_w) = 1/(\phi_e \underline{\phi})$  is independent of the worker productivity  $\phi_w$  by (13).

Since profits are supermodular in productivities, there will be PAM between workers and entrepreneurs: more productive entrepreneurs will hire more productive workers. The equilibrium assignment is then such that  $x\%$  of the labor demand coming from the least productive entrepreneurs is satisfied by  $x\%$  of the labor supply of the least productive workers (Sattinger, 1993; Legros and Newman, 2002). More formally, the matching condition between workers and entrepreneurs is based on the equality of differential increments of labor supply and labor demand. The increment of labor supply corresponds to the increment of higher-ability workers in the population:

$$dL^S = Lf(\phi_w)d\phi_w, \quad (17)$$

whereas the increment of labor demand corresponds to the increment of output produced by more productive firms:

$$dL^D = \frac{X(\phi_e)}{\phi_e\phi_w} f(\phi_e)d\phi_e. \quad (18)$$

Equating labor supply (17) and labor demand (18), we can find the assignment  $\phi_w = \phi_w(\phi_e)$  or workers to firms. By definition, the conditions  $\phi_w(\widehat{\phi}) = \underline{\phi}$  and  $\phi_w(\overline{\phi}) = \widehat{\phi}$  must hold on the boundaries of the productivity support.

## 2.5 Equilibrium

We now spell out the equilibrium conditions of our model. An equilibrium is such that: (i) all agents maximize utility; (ii) agents make optimal occupational choices; (iii) agents make optimal locational choices within the city; (iv) entrepreneurs maximize profits by setting optimal prices and hiring the optimal type of worker; (v) goods markets clear; and (vi) the labor market clears for all types of workers. Formally, an equilibrium is determined from the following conditions:

- Occupational choice:  $\pi^*(\widehat{\phi}) = w(\widehat{\phi})$
- Labor market clearing:  $L^S(\phi_w) = L^D(\phi_e(\phi_w))$  for all  $\phi_w \in [\underline{\phi}, \widehat{\phi}]$ .
- Utility maximization:  $\frac{u'(x(\nu))}{\lambda(\phi)} = p(\nu)$  for all  $\nu \in \Omega_e$  and  $\phi \in [\underline{\phi}, \overline{\phi}]$
- Profit maximization:  $p(\phi_e, \phi_w) = \frac{1}{\phi_e\phi} \frac{1}{1-r(\phi_e, \phi_w)}$  for all  $\phi_e \in \Omega_e$
- Budget constraints:  $\int_{\Omega_e} p(\nu)x(\nu)dF(\nu) + R(l) = I(\phi)e^{-\tau l}$  for all  $\phi \in [\underline{\phi}, \overline{\phi}]$

We solve these conditions for the cutoff  $\widehat{\phi}$ , the assignment of workers to firms  $\phi_w = \phi_w(\phi_e)$ , individual demands  $x(\nu)$ , prices  $p(\phi_e, \phi_w)$ , and the distribution of Lagrange multipliers  $\lambda(\phi)$ . As discussed in Section 2.1, when the Lagrange multipliers can be aggregated into a ‘generalized multiplier’  $\Lambda$ , the problem is substantially simplified since we can just solve for that multiplier instead of for the whole distribution. In all the examples we develop later in Section 4, there exists a generalized multiplier  $\Lambda$  that we can determine uniquely from the equilibrium conditions above.

## 2.6 Inequality measures

One of our key objectives is to investigate how city size,  $L$ , and the underlying productivity distribution,  $F$ , map into city-wide income distributions, especially consumers' disposable incomes. As is well known, there are many ways to look at income inequality. Standard measures include the Gini coefficient, entropy-type measures, the coefficient of variation of incomes, and the 90-10 income quantile ratio.

In what follows, we will look at inequality both at the city level, and at the level of individual matches. In the case of individual matches, the easiest measure to look at is the quantile ratio. Let  $\mathcal{H}$  denote the set of 'high income' agents (e.g., at the 90th percentile of the distribution), and let  $\mathcal{L}$  denote the set of 'low income' agents (e.g., at the 10th percentile of the distribution). Two intuitive measures of inequality are then given by:

$$\rho_1(\phi_h, \phi_l) \equiv \frac{\pi(\phi_h)}{\pi(\phi_l)}, \quad \rho_2(\phi_h, \phi_l) \equiv \frac{\pi(\phi_h)}{w(\phi_l, \phi_h(\phi_l))} \quad (19)$$

where  $h \in \mathcal{H}$  and  $l \in \mathcal{L}$  denote 'high' and 'low' income agents, respectively. The first measure looks at income inequality between entrepreneurs. The second measure looks at the split between wages and profits at the level of the firm.

At the city level, we will look (numerically) at the coefficient of variation in disposable incomes for all agents:

$$\text{CV}(\text{DI}) = \frac{\sigma(\text{DI})}{\overline{\text{DI}}},$$

where DI is defined as in (7), and where  $\sigma(\text{DI})$  and  $\overline{\text{DI}}$  are the standard deviation and the average of the distribution of disposable incomes, respectively. Last, we can also look at the split of income into wages and profits at the level of the city:

$$\rho_3(\phi_h, \phi_l) \equiv \frac{\int_{\mathcal{H}} \pi(\phi_h) dF(\phi_h)}{\int_{\mathcal{L}} w(\phi_l) dF(\phi_l)}.$$

## 3 Incomes and location

Distribution of types of agents  $F(\phi) : [\underline{\phi}; \bar{\phi}] \mapsto [0; 1]$  and distribution of land  $G(l) : [0; \bar{l}] \mapsto [0; 1]$  satisfy following properties:  $\frac{d^2 F}{d\phi^2} < 0$ ,  $\frac{d^2 G}{dl^2} > 0$

Then location assignment  $\phi(l)$  is generated by the rule

$$F(\phi(l)) = 1 - G(l), \quad \phi(\bar{l}) = \underline{\phi}, \quad \phi(0) = \bar{\phi}$$

yields the properties:  $\phi'(l) < 0$

*Proof.*  $\frac{dF}{d\phi} \phi'(l) = -\frac{dG}{dl}$ ,  $\phi'(l) = -\frac{\frac{dG}{dl}}{\frac{dF}{d\phi}} < 0$ , because  $\frac{dF}{d\phi} > 0$  and  $\frac{dG}{dl} > 0$  as a probability distribution functions.  $\square$

Let's consider preferences generating following properties:

1. wage and profit are increasing functions of agent's type:  $w'(\phi) > 0$ ,  $\pi'(\phi) > 0$
2. profit is steeper than wage at the cutoff point:  $\pi'(\hat{\phi}) > w'(\hat{\phi})$
3. measure of concavity of profit greater than wage's one:  $\mathcal{E}\pi'(\phi) = \frac{\pi''(\phi)\phi}{\pi'(\phi)} > \frac{w''(\phi)\phi}{w'(\phi)} = \mathcal{E}w'(\hat{\phi})$

Then income distribution of agents along city locations  $I(l)$  satisfies properties:

- $I'(l) < 0$
- $I_w(\hat{l}) = I_E(\hat{l})$
- $I'_E(\hat{l}) > I'_w(\hat{l})$
- $\mathcal{E}I'_E(l) > \mathcal{E}I'_w(l)$

*Proof.*  $I_w(\hat{l}) = w(\hat{\phi}) = \pi(\hat{\phi}) = I_E(\hat{l})$

$$I'(l) = \frac{dI}{d\phi}\phi'(l) < 0$$

$$I'_E(\hat{l}) = \pi'(\hat{\phi})\phi'(\hat{l}) > w'(\hat{\phi})\phi'(\hat{l}) = I'_w(\hat{l})$$

$$\mathcal{E}I'_E(l) = \mathcal{E}\pi'(\phi)\mathcal{E}\phi(l) > \mathcal{E}w'(\phi)\mathcal{E}\phi(l) = \mathcal{E}I'_w(l) \quad \square$$

To derive sharper results, we now look at incomes and locational choices given some parametrization of the productivity distribution and the distribution of land within the city.

The land rent  $r(l)$  is obtained from the locational equilibrium in the mono-centric city, which assigns types to locations ( $\phi = \phi(l)$ ). Note that under the fixed lot size assumption, the location assignment does not depend on agents' preferences. Hence, with a uniform distribution for  $\phi$ , condition (8) reduces to

$$\frac{\phi^* - \underline{\phi}}{\bar{\phi} - \underline{\phi}} = \frac{l(\underline{\phi}) - l(\phi^*)}{\bar{l} - 0}, \quad \forall \phi^* \in [\underline{\phi}, \bar{\phi}],$$

with boundaries constraints:  $l(\bar{\phi}) = 0$ ,  $l(\underline{\phi}) = \bar{l}$ . It follows that

$$\frac{\phi^* - \underline{\phi}}{\bar{\phi} - \underline{\phi}} = \frac{\bar{l} - l(\phi^*)}{\bar{l}}, \quad \forall \phi^* \in [\underline{\phi}, \bar{\phi}],$$

so that we finally get

$$l(\phi) = \frac{\bar{\phi} - \phi}{\bar{\phi} - \underline{\phi}}\bar{l} \quad \text{or, equivalently,} \quad \phi(l) = \bar{\phi} - \frac{l}{\bar{l}}(\bar{\phi} - \underline{\phi}). \quad (20)$$

Note also that the income of workers is independent of the preferences. Hence, since  $I(\phi_w) = w(\phi_w) = \frac{\phi_w}{\underline{\phi}}$ , we can make use of the location assignment to obtain income as a function of location as follows:

$$I(l) = 1 - \frac{l}{\bar{l}} \left( \frac{\bar{\phi}}{\underline{\phi}} - 1 \right), \quad \text{and therefore} \quad r'(l) = \left[ -\tau + \frac{\tau l}{\bar{l} \underline{\phi}} (\bar{\phi} - \underline{\phi}) \right] e^{-\tau l}$$

Integrating and using the normalization  $r(\bar{l}) = 0$ , the rent gradient on the domain of workers – i.e., on  $l \in [l(\hat{\phi}), l(\underline{\phi})]$  – is given by

$$r(l) = e^{-\tau l} \left( 1 - \frac{l}{\bar{l}} \right) \left( \frac{\bar{\phi}}{\underline{\phi}} - 1 \right) + \frac{1}{\tau \bar{l}} \left( \frac{\bar{\phi}}{\underline{\phi}} - 1 \right) \left( e^{-\tau \bar{l}} - e^{-\tau l} \right) - \left( e^{-\tau \bar{l}} - e^{-\tau l} \right). \quad (21)$$

For entrepreneurs, the land rent schedule is more complex to determine since it depends on the endogenously determined cutoff  $\hat{\phi}$ . Yet, observe that the full rent schedule will be continuously differentiable since

$$r'(l(\hat{\phi})) = -\tau e^{-\tau l} I(\hat{\phi})$$

applies to both workers and entrepreneurs by definition of  $\hat{\phi}$ :  $\pi(\hat{\phi}) = I_e(\hat{\phi}) = \hat{\phi}/\underline{\phi} = I_w(\hat{\phi})$ , so that  $r'_e(l(\hat{\phi})) = r'_w(l(\hat{\phi}))$ .

## 4 Specific examples and numerical illustrations

We now derive sharper results – including comparative statics – by focusing more explicitly on specific cases. We first develop the ‘standard’ constant-elasticity-of-substitution (CES) case, where city size has no impact on occupational selection. We then also work through the logarithmic case where city size has an impact on occupational selection. To ease the exposition, we first consider a uniform distribution of types. We relax this condition later on and consider arbitrary distributions.

Note that all the cases we discuss allow for a decomposition of demand using a ‘generalized’ lambda,  $\Lambda$ . The advantage of these cases is that we do not have to solve for the *distribution* of the multipliers  $\lambda(\phi)$ : knowing one aggregate statistic is enough.

### 4.1 CES preferences

Assume that  $F(\phi) = \frac{\phi - \underline{\phi}}{\bar{\phi} - \underline{\phi}}$  and that  $u(x) = x^\rho$ , with  $0 < \rho < 1$ . The first-order conditions of consumers imply that

$$\frac{\rho [x(\phi_e)]^{\rho-1}}{\lambda(\phi_w)} = p(\phi_e), \quad (22)$$

whereas the first-order conditions of producers imply that

$$r(\phi_e, \phi_w) = \frac{p(\phi_e) - \frac{w(\phi_e, \phi_w)}{\phi_e \phi_w}}{p(\phi_e)} = 1 - \rho, \quad (23)$$

because of the constant elasticity of demand. Hence, we have

$$p(\phi_e) = \frac{w(\phi_e)}{\rho \phi_e \phi_w} \quad \text{and} \quad x(\phi_e | \phi) = \left[ \frac{\lambda(\phi) w(\phi_e)}{\rho^2 \phi_e \phi_w} \right]^{\frac{1}{\rho-1}}.$$

Clearly, demand is multiplicatively quasi-separable (Behrens and Murata, 2007), so that the aggregate output of a type- $\phi_e$  firm is given by:

$$X(\phi_e) = \left[ \frac{w(\phi_e, \phi_w)}{\rho^2 \phi_e \phi_w} \right]^{\frac{1}{\rho-1}} \int_{\Omega} [\lambda(\phi)]^{\frac{1}{\rho-1}} dF(\phi)$$

The maximized profit of a type- $\phi_e$  firm – conditional on hiring type- $\phi_w$  workers – is thus

$$\pi^*(\phi_e | \phi_w) = \left( \frac{1 - \rho}{\rho} \right) \frac{w(\phi_e, \phi_w)}{\phi_e \phi_w} \left[ \frac{w(\phi_e, \phi_w)}{\rho^2 \phi_e \phi_w} \right]^{\frac{1}{\rho-1}} \int_{\Omega} [\lambda(\phi)]^{\frac{1}{\rho-1}} dF(\phi),$$

which can be more conveniently expressed as

$$\pi^*(\phi_e | \phi_w) = (1 - \rho) \rho^{\frac{1+\rho}{1-\rho}} \left[ \frac{w(\phi_e, \phi_w)}{\phi_e \phi_w} \right]^{\frac{\rho}{\rho-1}} \Lambda^{\frac{1}{\rho-1}}, \quad \Lambda = \left[ \int_{\Omega} [\lambda(\phi)]^{\frac{1}{\rho-1}} dF(\phi) \right]^{\rho-1},$$

where  $\Lambda$  is an aggregate that is taken as given by each individual firm. Conditional on a given wage rate, it can readily be verified that  $(\partial^2 \pi^* / (\partial \phi_e \partial \phi_w)) > 0$ , i.e., profits are supermodular in the two productivity parameters. Hence, there will be positive assortative matching (PAM) between firms and workers.

Using the results of section 2.3, we know that  $w(\phi_w) = \phi_w / \underline{\phi}$ . Hence, the profit of a type- $\phi_e$  firm is:

$$\pi(\phi_e | \phi_w) = (1 - \rho) \rho^{\frac{1+\rho}{1-\rho}} \left[ \underline{\phi} \phi_e \right]^{\frac{\rho}{1-\rho}} \Lambda^{\frac{1}{\rho-1}}$$

From the cutoff condition  $\pi^*(\widehat{\phi}) = \widehat{\phi} / \underline{\phi}$ , we then have

$$(1 - \rho) \rho^{\frac{1+\rho}{1-\rho}} \left[ \underline{\phi} \widehat{\phi} \right]^{\frac{\rho}{1-\rho}} \Lambda^{\frac{1}{\rho-1}} = \frac{\widehat{\phi}}{\underline{\phi}} \iff (1 - \rho) \rho^{\frac{1+\rho}{1-\rho}} \left[ \underline{\phi} \widehat{\phi} \right]^{\frac{\rho}{1-\rho}} \Lambda^{\frac{1}{\rho-1}} = \frac{\widehat{\phi}}{\underline{\phi}},$$

which can be solved for the ‘aggregate Lagrange multiplier’  $\Lambda$  as follows:

$$\Lambda^{\frac{1}{\rho-1}} = (1 - \rho) \rho^{\frac{1+\rho}{1-\rho}} \frac{\widehat{\phi}}{\underline{\phi}} \left[ \underline{\phi} \widehat{\phi} \right]^{\frac{\rho}{1-\rho}} \quad (24)$$

Substituting (24) back into profits finally yields:

$$\pi(\phi_e) = \left[ \frac{\phi_e}{\widehat{\phi}} \right]^{\frac{\rho}{1-\rho}} \frac{\widehat{\phi}}{\underline{\phi}}.$$

The foregoing expression highlights that the curvature of the rent function is different in the part occupied by workers from that in the part occupied by entrepreneurs. An appropriate measure of curvature of the rent function is the elasticity of its derivative. Calculating it, we have:  $\mathcal{E}r'(l) = -\tau l + \mathcal{E}I(l)$ , where  $\mathcal{E}I(l) = \mathcal{E}I(\phi) \cdot \mathcal{E}\phi(l)$ . At the location of the cutoff  $\hat{\phi}$ , the difference of curvatures is due to the difference of elasticities of income of workers and entrepreneurs:  $\mathcal{E}\pi(\phi) = \frac{\rho}{1-\rho}$ , whereas  $\mathcal{E}w(\phi) = 1$ . Thus, the curvature of the rent function for entrepreneurs is greater than that for workers when  $\rho > 1/2$ , which we assume to be the case in what follows.

Note also that another conditions we have to pay attention to is an inequality of derivatives between profit and reservation wage for the cutoff-type agent. Namely, the following condition must hold true:

$$\pi'(\hat{\phi}) \geq w'(\hat{\phi})$$

In other words we have following constraint:

$$\frac{\rho}{1-\rho} \frac{\hat{\phi}}{\underline{\phi}} \geq \frac{1}{\underline{\phi}}, \quad \text{which means that} \quad \hat{\phi} \geq 1 - \frac{1}{\rho}.$$

In the plausible case where  $\rho \geq 1/2$ , this means that  $\hat{\phi}$  must exceed some value less than one. Recalling that  $\hat{\phi} \geq \underline{\phi}$ , and making the additional innocuous assumption that  $\underline{\phi} \geq 1$ , we get the inequality  $\hat{\phi} \geq 1 - \frac{1}{\rho}$ . It holds by inspection for any exogeneous  $\rho$  and any endogeneous  $\hat{\phi}$ .

Turning to the assignment of workers to firms, we have

$$dL^D = \frac{1}{\phi_e \phi_w} \left[ \frac{w(\phi_e, \phi_w)}{\rho^2 \phi_e \phi_w} \right]^{\frac{1}{\rho-1}} \Lambda^{\frac{1}{\rho-1}} d\phi_e = \frac{1}{\rho^{\frac{2}{\rho-1}} \phi_w \phi_e^{\frac{1}{\rho-1}} \phi_e^{\frac{\rho}{\rho-1}}} \Lambda^{\frac{1}{\rho-1}} d\phi_e.$$

Equating labor demand and supply,  $dL^D = dL^S$ , we have:

$$\phi_w d\phi_w = \rho^{\frac{2}{1-\rho}} \underline{\phi}^{\frac{1}{1-\rho}} \phi_e^{\frac{\rho}{1-\rho}} \Lambda^{\frac{1}{\rho-1}} d\phi_e,$$

which can be integrated to yield

$$\frac{\phi_w^2}{2} = (1-\rho) \rho^{\frac{2}{1-\rho}} \underline{\phi}^{\frac{1}{1-\rho}} \phi_e^{\frac{\rho}{1-\rho}} \Lambda^{\frac{1}{\rho-1}} + K.$$

Using the boundary condition  $\phi_w(\bar{\phi}) = \hat{\phi}$  we can obtain the assignment:

$$\frac{\phi_w^2}{2} = \frac{\hat{\phi}^2}{2} + \Lambda^{\frac{1}{\rho-1}} (1-\rho) \rho^{\frac{2}{1-\rho}} \underline{\phi}^{\frac{1}{1-\rho}} \left( \phi_e^{\frac{\rho}{1-\rho}} - \bar{\phi}^{\frac{\rho}{1-\rho}} \right).$$

Using the second boundary condition  $\phi_w(\hat{\phi}) = \underline{\phi}$  we finally get another expression for  $\Lambda$ :

$$\Lambda^{\frac{1}{\rho-1}} = \frac{1}{2} \frac{\hat{\phi}^2 - \underline{\phi}^2}{(1-\rho) \rho^{\frac{2}{1-\rho}} \underline{\phi}^{\frac{1}{1-\rho}} \left( \bar{\phi}^{\frac{\rho}{1-\rho}} - \hat{\phi}^{\frac{\rho}{1-\rho}} \right)}$$

The equilibrium cutoff  $\hat{\phi}$  is the solution to a system of two equations:

$$\begin{cases} \Lambda^{\frac{1}{\rho-1}} = \frac{\hat{\phi}^{\frac{1-2\rho}{1-\rho}}}{(1-\rho)\rho^{\frac{1+\rho}{1-\rho}}\phi^{\frac{1}{1-\rho}}} \\ \Lambda^{\frac{1}{\rho-1}} = \frac{1}{2} \frac{\hat{\phi}^2 - \phi^2}{(1-\rho)\rho^{\frac{2}{1-\rho}}\phi^{\frac{1}{1-\rho}}(\bar{\phi}^{\frac{1}{1-\rho}} - \hat{\phi}^{\frac{1}{1-\rho}})} \end{cases} \quad (25)$$

The first equation describes a curve that takes positive finite values on  $[\underline{\phi}, \bar{\phi}]$ . It is decreasing when  $\rho < 1/2$ , and increasing when  $\rho > 1/2$ . The second equation describes an increasing curve, going from zero to infinity as  $\hat{\phi}$  changes from  $\underline{\phi}$  to  $\bar{\phi}$ . Thus, an equilibrium exists and is unique.

[Equilibrium under CES] Assume that skills are uniformly distributed, and that the preferences are of the CES-type. Then a unique equilibrium cutoff  $\hat{\phi} \in (\underline{\phi}, \bar{\phi})$  exists.

**Numerical illustrations.** Let us illustrate numerically the model. Let  $\rho = 0.7$ ,  $L = 10$ ,  $\underline{\phi} = 1$  and  $\bar{\phi} = 2$ . This yields a cutoff given by  $\hat{\phi} = 1.32$ , associated with the distance  $\hat{l} = l(\hat{\phi}) = 6.84$ . The coefficient of variation of disposable income,  $CV_{DI} = \sigma_{DI}/\mu_{DI} = 2.04$ .

Now consider a doubling of the city size  $L = 20$ . The cutoff  $\hat{\phi}$  does not change, but the cutoff location naturally ‘doubles’:  $\hat{l} = 13.69$ . Despite the unchanged selection cutoff, inequality as measured by disposable income increases a lot:  $CV_{DI} = \sigma_{DI}/\mu_{DI} = 3.17$ .

## 4.2 Logarithmic preferences

One of the particular properties of the CES model is that the selection cutoff is independent of the size of the city,  $L$  (see also Behrens, Duranton, and Robert-Nicoud, 2012). Assume now that  $F(\phi) = \frac{\phi - \underline{\phi}}{\bar{\phi} - \underline{\phi}}$  but that  $u(x) = \ln(x + 1)$ . The first-order conditions of consumers imply that

$$x(\phi_e) = \frac{1}{p(\phi_e)\lambda(\phi)} - 1$$

so that aggregate demand can be expressed as follows:

$$X(\phi_e) = L \left[ \frac{\Lambda}{p(\phi_e)} - 1 \right], \quad \text{with} \quad \Lambda = \int_{\Omega} \lambda^{-1}(\phi) dF(\phi).$$

Straightforward profit maximization then yields the profit-maximizing price

$$p^*(\phi_e | \phi_w) = \left[ \frac{w(\phi_e, \phi_w)\Lambda}{\phi_e\phi_w} \right]^{\frac{1}{2}},$$

and maximum profits:

$$\pi^*(\phi_e | \phi_w) = L \frac{w(\phi_e, \phi_w)}{\phi_e\phi_w} \left\{ \Lambda^{\frac{1}{2}} \left[ \frac{w(\phi_e, \phi_w)}{\phi_e\phi_w} \right]^{-\frac{1}{2}} - \frac{w(\phi_e, \phi_w)}{\phi_e\phi_w} \right\} \left\{ \Lambda^{\frac{1}{2}} \left[ \frac{w(\phi_e, \phi_w)}{\phi_e\phi_w} \right]^{-\frac{1}{2}} - 1 \right\}$$

As before, wages are such that  $w(\phi_w) = \phi_w/\underline{\phi}$ , so that the maximized profit is given by

$$\pi^*(\phi_e) = L \frac{1}{\phi_e \underline{\phi}} \left\{ \Lambda^{\frac{1}{2}} [\phi_e \underline{\phi}]^{\frac{1}{2}} - [\phi_e \underline{\phi}]^{-1} \right\} \left\{ \Lambda^{\frac{1}{2}} [\phi_e \underline{\phi}]^{\frac{1}{2}} - 1 \right\}$$

The cutoff condition  $\pi^*(\hat{\phi}) = w(\hat{\phi}) = \hat{\phi}/\underline{\phi}$  can be solved for  $\Lambda$  to yield a first equilibrium relationship:

$$\Lambda = \frac{L + L(\hat{\phi}\underline{\phi})^2 + 2\hat{\phi}^2(\hat{\phi}\underline{\phi})^2 + \sqrt{L}(1 + \hat{\phi}\underline{\phi})\sqrt{4\hat{\phi}^4\underline{\phi}^2 + L(\hat{\phi}\underline{\phi} - 1)^2}}{2L\hat{\phi}^3\underline{\phi}^3}. \quad (26)$$

Note that, contrary to the CES case, this relationship depends in a non-proportional way on city size  $L$ .

Equating labor supply and demand yields

$$dL^s = \frac{1}{\hat{\phi} - \underline{\phi}} d\phi_w = dL^d = \frac{1}{\hat{\phi} - \underline{\phi}} \frac{L}{\phi_e \phi_w} \left[ \sqrt{\Lambda \phi_e \underline{\phi}} - 1 \right] d\phi_e \Rightarrow \frac{d\phi_w}{d\phi_e} = \frac{L}{\phi_e \phi_w} \left[ \sqrt{\Lambda \phi_e \underline{\phi}} - 1 \right].$$

This differential equation can be readily solved for the assignment  $\phi_w(\phi_e)$ . Using the boundary condition  $\phi_w(\hat{\phi}) = \underline{\phi}$  allows to pin down the constant, whereas the second boundary condition  $\phi_w(\bar{\phi}) = \hat{\phi}$  gives a second condition linking the cutoff  $\hat{\phi}$  and the aggregate Lagrange multiplier  $\Lambda$ . Solving that relationship for  $\Lambda$  yields:

$$\Lambda = \frac{\left[ \hat{\phi}^2 - \underline{\phi}^2 + 2L(\underline{\phi} - \bar{\phi}) \ln \hat{\phi} - 2L(\underline{\phi} - \bar{\phi}) \ln \underline{\phi} \right]^2}{16L^2 \underline{\phi} (\underline{\phi} - \bar{\phi})^2 (\sqrt{\hat{\phi}} - \sqrt{\bar{\phi}})^2} \quad (27)$$

Equations (26) and (27) can be solved for  $\hat{\phi}$ , which depends on city size  $L$ .

## 5 Conclusions

## References

- [1] Beckmann, Martin J. (1969) On the Distribution of Urban Rent and Residential Density, *Journal of Economic Theory* 1(1), 60–67.
- [2] Behrens, Kristian (2013) Geography and Trade: Unweaving Textile Clusters in Canada. Univ. of Quebec at Montreal, processed.
- [3] Behrens, Kristian, and Théophile Bougna (2013) An anatomy of the geographical concentration of Canadian manufacturing industries. CRIPÉE Discussion Paper #13–27.
- [4] Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Südekum (2011) Spatial Frictions. CEPR Discussion Paper #8572, Center for Economic Policy Research.

- [5] Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud (2013) Productive cities: Agglomeration, selection, and sorting. Conditionally accepted at the *Journal of Political Economy*.
- [6] Behrens, Kristian, and Frédéric Robert-Nicoud (2013) Survival of the fittest in cities: Urbanisation and inequality, *Economic Journal*, forthcoming.
- [7] Bernard, Andrew B., and J. Bradford Jensen (1995) Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987, *Brookings Papers on Economic Activity, Microeconomics*, 67–119
- [8] Bernard, Andrew B., and J. Bradford Jensen (2007) Firm structure, multinationals, and manufacturing plant deaths, *Review of Economics and Statistics* 89(2), 193–204.
- [9] Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux (2012) Sorting and local wage and skill distributions in France, *Regional Science and Urban Economics* 42, 913–930.
- [10] Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon (2012) The Costs of Agglomeration: Land Prices in French Cities. Wharton, Univ. of Pennsylvania, processed.
- [11] Davis, Donald R., and Jonathan I. Dingle (2013) A spatial knowledge economy. Columbia Univ., processed.
- [12] D’Costa, Sabine (2010) Trade liberalization and the geographical concentration of industries: Evidence from Hungarian manufacturing industries. PhD Dissertation, London School of Economics.
- [13] Duranton, Gilles, and Henry G. Overman (2005) Testing for Localisation Using Micro-Geographic Data, *Review of Economic Studies* 72(4), 1077–1106.
- [14] Duranton, Gilles, and Henry G. Overman (2008) Testing for Localisation Using Micro-Geographic Data, *Journal of Regional Science* 48(1), 213–243.
- [15] Ellison, Glenn D., Edward L. Glaeser, and William R. Kerr (2010) What causes industry agglomeration? Evidence from coagglomeration patterns, *American Economic Review* 100(3), 1195–1213.
- [16] Fujita, Masahisa, and Jacques-François Thisse (2006) Globalization and the Evolution of the Supply Chain: Who Gains and Who Loses?, *International Economic Review* 47(3), 811–836.
- [17] Hecker, Daniel E. (2005) High-technology employment: a NAICS-based update, *Monthly Labor Review* 128(7), 57–72.
- [18] Helsley, Robert W., and William C. Strange (2012) Co-agglomeration and the size and composition of clusters. Univ. of California, Berkeley; and Rotmann School, Univ. of Toronto, processed.

- [19] Holmes, Thomas J. (1999) Localization of industry and vertical disintegration, *Review of Economics and Statistics* 81(2), 314–325.
- [20] Koenig, Pamina (2009) Agglomeration and the export decisions of French firms, *Journal of Urban Economics* 66, 186–195.
- [21] Koenig, Pamina, Florian Mayneris and Sandra Poncet (2010) Local export spillovers in France, *European Economic Review* 54, 622–641.
- [22] Krugman, Paul R., and Anthony J. Venables (1995) Globalization and the Inequality of Nations, *The Quarterly Journal of Economics* 110(4), 857–880.
- [23] Lucas, Robert E. (1978) On the Size Distribution of Business Firms, *Bell Journal of Economics* 9(2), 508–523.
- [24] Lucas, Robert E., and Esteban Rossi-Hansberg (2002) On the Internal Structure of Cities, *Econometrica* 70(4), 1445–1476.
- [25] Melitz, Marc J. (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity, *Econometrica* 71, 1695–1725.
- [26] Melitz, Marc J., and Gianmarco I.P. Ottaviano (2008) Market size, trade, and productivity, *Review of Economic Studies* 75, 295–316.
- [27] Ševčík, Pavel (2012) Plant Size, Productivity, and the Efficiency of Corporate Diversification. Univ. of Quebec at Montreal, processed.