DISCUSSION PAPER SERIES

DP19016

HOW AGGREGATE UNCERTAINTY SHAPES THE SPATIAL ECONOMY

Peter Egger, Katharina Erhardt and Davide Suverato

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

HOW AGGREGATE UNCERTAINTY SHAPES THE SPATIAL ECONOMY

Peter Egger, Katharina Erhardt and Davide Suverato

Discussion Paper DP19016
Published 24 April 2024
Submitted 11 April 2024

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

HOW AGGREGATE UNCERTAINTY SHAPES THE SPATIAL ECONOMY

Abstract

This paper develops a dynamic spatial general equilibrium model of a multi-region multi-sector open economy in which heterogeneous agents choose optimally where to locate and work, by making forward-looking decisions under aggregate uncertainty about future realizations of economic fundamentals and idiosyncratic risk of aging. To address the role of aggregate uncertainty in interaction with spatial determinants of the equilibrium, we solve the system of individual dynamic optimal-control problems under rational expectations as a Mean Field Game, in discrete time, over the discrete state space, preserving the full non-linear structure of the problem. With a calibration for France, we demonstrate that households behave substantially differently between uncertainty and perfect foresight. By affecting the continuation value of jobs -- differently by location and age -- uncertainty alters the patterns of labor reallocation in transition as well as in the long run. The impact of uncertainty alone on individual lifetime welfare is negative on average, but it triggers heterogeneous welfare changes: substantial portions of the population lose much more than the average (are stuck in less attractive jobs) while some agents actually gain (since others did not reallocate). Consequently, the spatial distribution of economic activity deviates ceteris paribus systematically when agents make decisions under uncertainty compared to perfect foresight. In that sense, aggregate uncertainty per se shapes the spatial economy.

JEL Classification: F16, J62, R13

Keywords: Economic geography, Mean field games

Peter Egger - pegger@ethz.ch
ETH Zurich and CEPR

Katharina Erhardt - erhardt@dice.hhu.de
einrich-Heine-University Düsseldorf

Davide Suverato - dsuverato@ethz.ch
ETH Zurich

How Aggregate Uncertainty Shapes the Spatial Economy*

Peter H. Egger† Katharina Erhardt‡ Davide Suverato§

March 2024

Abstract

This paper develops a dynamic spatial general equilibrium model of a multi-region multi-sector open economy where heterogeneous agents choose optimally their job, making forward-looking decisions under aggregate uncertainty. We solve the system of individual dynamic optimal-control problems under rational expectations as a Mean Field Game, in discrete time and state space, preserving the full non-linear structure of the problem. With a calibration for France, we demonstrate that households behave substantially differently between uncertainty and perfect foresight. Uncertainty alters the patterns of labor reallocation in transition as well as in the long run.

JEL Codes: E24, F16, J62, R13

Keywords: Spatial dynamic general equilibrium, Uncertainty, Trade and labor market interactions, Mean field games.

*Acknowledgments: We are grateful to David Argente, Costas Arkolakis, Erhan Artuç, David Atkin, Andy Bernard, Adrien Bilal, Johannes Boehm, Lorenzo Caliendo, Paola Conconi, Bruno Conte, Arnaud Costinot, Rafael Dix-Carneiro, Amit Khandelwal, Sam Kortum, Jie Li, John McLaren, Isabelle Méjean, Giordano Mion, Fernando Parro, Frédéric Robert-Nicoud, Farid Toubal, Gianluca Violante, and Miaojie Yu for numerous comments on earlier versions of the paper. We are also thankful to the seminar participants at CESifo, ERWIT, MIT, Paris School of Economics, Stavanger, and Yale for their useful questions and insightful discussions. Finally, we are grateful to the math professors Francesca Da Lio and Herbert Egger for their insights and their availability to discuss recent advances in their fields with us. This paper is based on data from Eurostat (ERFT 2011-2017). The responsibility for all conclusions drawn from the data lies entirely with the authors.

†Corresponding author. ETH Zurich, CEPR, CESifo, GEP, WIFO, Address: ETH Zurich, LEE G101, Leonhardstrasse 21, 8092 Zurich, Switzerland; E-mail: pegger@ethz.ch

‡Heinrich-Heine-University Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), CESifo, Address: Universitätsstr. 1, 40225 Düsseldorf, Germany; Email: erhardt@dice.hhu.de;

§ETH Zurich, CAREFIN Bocconi University, Address: ETH Zurich, LEE G101, Leonhardstrasse 21, 8092 Zurich, Switzerland; E-mail: dsuverato@ethz.ch.

1 Introduction

Economists and policymakers alike are interested in understanding how the spatial allocation of economic activity is affected by changes of the fundamentals of an economy. These changes can result from trade-policy shocks (Caliendo et al., 2019; Dix-Carneiro et al., 2023), trends in robotization (Acemoglu and Restrepo, 2020), climate change (Bilal and Rossi-Hansberg, 2023), or the sudden emergence of geopolitical tensions (Baker et al., 2023). Spatial general equilibrium models are a workhorse tool for quantifying how changes of the determinants of local economic activity – such as productivity or market access – lead to a restructuring of the economic geography through spatial linkages. However, one aspect that has received little attention in spatial general-equilibrium models is that agents act under uncertainty about the actual realizations of economic fundamentals in response to shocks, and that the uncertain component of these fundamentals is itself an important determinant of reallocation decisions.

While choices about jobs and residence locations are in the limelight of spatial models, the future value of jobs or residence locations must often be predicted long before and without knowing to which extent expected benefits materialize – hence, decisions are taken under uncertainty. So far, we know relatively little about how the uncertainty about economic fundamentals affects the spatial equilibrium: the vast majority of existing frameworks solve for an equilibrium in which agents have perfect foresight on aggregate economic outcomes, thus, shutting down the role of uncertainty for tractability reasons.[1]

Only recently, the literature has been enriched by frameworks that can handle uncertainty in spatial dynamic models.[2] Essentially, these build on the idea of perturbations around a perfect-foresight path. This makes them flexible and efficient but less suited for a context of shocks that are aggregate, large, and, hence, discrete in nature (both over time and space), where the anchor to a perfect-foresight trajectory might rise concerns. This is the context in which, in this paper, we aim to provide some closure to the research question examining how uncertainty influences the spatial distribution of economic activity.

We offer two contributions to the literature. We illustrate a solution methodology that allows for solving dynamic spatial general-equilibrium models under aggregate uncertainty.

[1] For instance, in the seminal contributions of Artuç et al. (2010) and Dix-Carneiro (2014) the problem of spatial labor reallocation is stated under uncertainty, and the structural estimation strategy accounts for it, but perfect foresight is then assumed and imposed to obtain the simulated predictions of the model.

[2] To the best of our knowledge, three approaches have been proposed to accomplish this task, namely the ones by Kleinman et al. (2021), Bilal (2023), and Fan et al. (2023). In what follows, and specifically in Section 5, we discuss similarities and differences of the approach proposed here with this earlier work.

This methodology rests on three fundamental pillars to overcome the curse of dimensionality standard computational methods face: (*i*) we cast the model as a coupled discrete-state and discrete-time Mean Field Game (MFG), following Gomes et al. (2010); (*ii*) we solve for the grid-valued counterpart of the MFG, extending the continuous-time approach by Achdou and Capuzzo-Dolcetta (2010) to discrete time and discrete state; (*iii*) we use barycentric piece-wise linear interpolation to write the MFG as a contraction on the discretized state space, building on the work by Debrabant and Jakobsen (2013).

Consequently, thanks to this novel approach, the key policy contribution of the present work is to highlight the role uncertainty plays as a determinant of the spatial allocation of economic activity. We show that uncertainty is not – only – a driver for freezing investment decisions (including mobility and specialization), as shown in the macro literature.[3] The spatial distribution of economic activity deviates systematically when agents make decisions under uncertainty compared to what would emerge under perfect foresight. Through this channel, uncertainty may generate a source of comparative (dis)advantage across jobs.

The spatial general-equilibrium model has the following features. (i) Each location hosts several sectors where workers in different occupations are employed. We will refer to a region-sector-occupation tuple as a 'job'. (ii) Given the allocation of labor at a given point in time, the within-period equilibrium is based on the model of Eaton and Kortum (2002) in its extension to multiple sectors by Caliendo and Parro (2015) and Caliendo et al. (2019). Production uses labor, locally fixed structures and materials that can be costly sourced from different places. (iii) Labor is mobile and can move across space at the end of each period subject to moving costs. (iv) Individuals move across space because of idiosyncratic preferences across jobs, different economic fundamentals across jobs, and different option values of being in particular jobs. Uncertainty about future real income associated with a job roots in the volatility of total factor productivity (TFP) which is specific to the production in a region and sector, as in Caliendo et al. (2018).[4]

This design is suited for a counterfactual analysis in which we assess the quantitative impact of uncertainty. Specifically, we compare the allocation of individuals across jobs un-

[3] The literature on real business cycle shows that an increase in volatility of returns freezes hiring and investment, Bloom (2009), which, in turns, triggers a drop in output (Fernandez-Villaverde et al., 2011, Bloom et al., 2018).

[4] While the way we look at uncertainty is common in macro and quantitative trade literature, there is a complementary view that focuses on the uncertainty of trade policy; see the two recent and insightful contributions by Pierce and Schott (2016) and Handley and Limao (2017). Furthermore, the role of uncertainty channeled via international trade has stimulated a large body on reduced-form studies; see Krishna et al. (2012) and Krishna and Senses (2014).

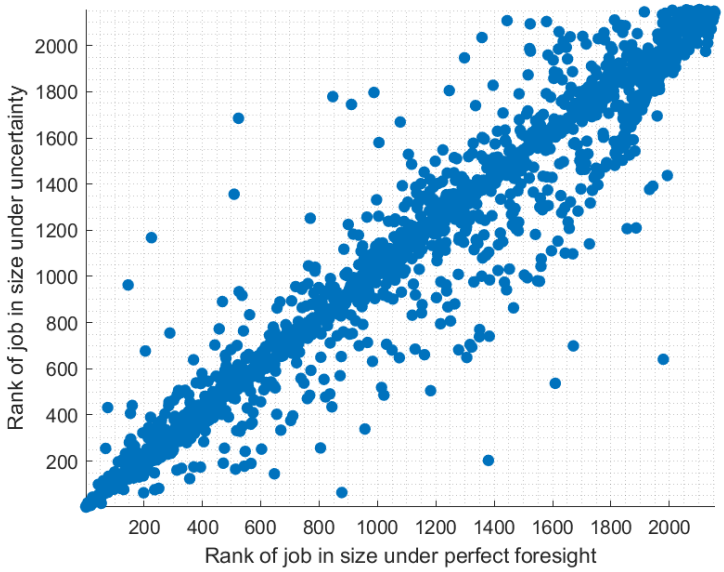


Figure 1.1: Ranking of jobs under uncertainty and perfect foresight

der uncertainty with the one under perfect foresight (certainty), while fixing all economic fundamentals in the two scenarios to be the same between these regimes, including actual TFP levels. To illustrate this point, we conduct a quantification of the model based on France and solve for the predicted distribution of labor over 2,156 job types.[5] Figure 1.1 plots the ranking of jobs by size (i.e., fraction of the workforce employed in a job) under uncertainty versus under perfect foresight resulting from this quantification for France, given the same fundamentals. While the mass of jobs on the diagonal confirms that the deterministic structure of the economy matters, the off-diagonal mass points to the differential impact of aggregate uncertainty: many jobs gain and others lose substantially in terms of employment size under uncertainty relative to perfect foresight.

What matters for the ranking of jobs in terms of their attractiveness is not uncertainty in absolute terms – any agent dislikes uncertainty, here – but in relative terms. Some jobs become more attractive compared to the perfect-foresight, as long as the uncertainty about the future value in interaction with economic fundamentals is considerably less important than elsewhere. Hence, although uncertainty is a cost on average, the value of jobs does not necessarily fall everywhere. The systematically different spatial allocation of labor across jobs under uncertainty leads to welfare losses for some and to welfare gains for others. Uncertainty-to-perfect-foresight welfare differences range from -5.46% to $+3.83\%$ among the

[5] Sections 6 and 7 provide a detailed discussion of the quantitative analysis.

older workforce (above 40) and from −17.03% to +4.37% among the young workforce. This heterogeneity materializes with an identical realization of economic fundamentals, hence, purely because of aggregate uncertainty.

Three bodies of work relate particularly closely to the problem addressed in this paper. First, within the quantitative trade and regional literature, our paper directly speaks to a recent line of research accounting for dynamic adjustments (see Desmet et al., 2018; Caliendo et al., 2019; Allen and Donaldson, 2020; Bilal and Rossi-Hansberg, 2021, Dix-Carneiro et al., 2023). However, prior work in this field did not and was not designed to address how global uncertainty affects the spatial distribution of economic agents and economic activity.[6] As already discussed, few recent approaches accommodate uncertainty, but as a perturbation of the approximated model around the steady state of the economy (such as in Kleinman et al., 2021 and Bilal, 2023), or around a perfect-foresight path (Fan et al., 2023). In comparison, our analysis is dynamic and preserves the full non-linear structure of the problem, while not relying on a steady state or perfect foresight. Thus, it is suited for addressing reallocation patterns that are discrete in nature and triggered by large, aggregate shocks.

Second, our work builds on the quantitative structural literature on labor-market adjustments following trade shocks (see Artuç et al., 2010; Dix-Carneiro and Kovak, 2017; Dix-Carneiro et al., 2021 and the review on local labor market adjustment by Moretti, 2011). Related work considers agents with perfect foresight about aggregate shocks. Artuç et al. (2010) develop a rational-expectations model of dynamic labor adjustment, then extended in Artuç and McLaren (2015) and widely adopted in the aforementioned literature. However, predictions of the approach are based on simulation of a perfect-foresight path of adjustment. The latter cannot offer a quantification of the importance of uncertainty. Dix-Carneiro and Kovak (2017) and Dix-Carneiro et al. (2021) carefully model switching costs, sector- and occupation-specific human capital, local amenities, and search frictions in labor markets, and agents eventually face uncertainty about idiosyncratic shocks, though not large, aggregate shocks and uncertainty.

A third line of research investigates the determinants of spatial allocation of economic activity, in interaction with geography and Ricardian comparative advantage; see the studies summarized in Redding and Rossi-Hansberg (2017). We join this literature by building on

[6] In Desmet et al. (2018) location choices do not depend on future economic conditions. In Caliendo et al. (2019) and Dix-Carneiro et al., 2023 agents are assumed to be perfectly foresighted. In Allen and Donaldson (2020) the driving force of reallocation (together with technology) is the accumulation of amenities. And in Bilal and Rossi-Hansberg (2021) agents' location decisions respond to income shocks which are not channeled through spatial linkages.

the modern quantitative, multi-country, multi-sector general-equilibrium framework in trade (see Eaton and Kortum, 2002; Dekle et al., 2007; Alvarez and Lucas, 2007; Caliendo and Parro, 2015) and regional economics where commuting or migration and a deeper understanding of spatial linkages are considered (see, e.g., Monte et al., 2018; Adao et al., 2020). Despite their usefulness for counterfactual analysis, these frameworks are static, thus, they are not suited for addressing the dynamics intrinsically rooted in forward-looking job and location decisions under aggregate uncertainty.

Finally, the paper is complementary to the macroeconomic literature on heterogeneous-agent models; see the review by Heathcote et al. (2009). In particular, we relate to frameworks that are based on the seminal contributions of Krusell and Smith (1998) and Reiter (2009), including the recent developments that build on MFG (Achdou et al., 2021; Bilal, 2023), on sequence-space Jacobian methods (Auclert et al., 2021), or on machine learning (Han et al., 2022). We share the research interest with this literature in the effect of uncertainty, here rooted in TFP-volatility shocks, on welfare. In the macroeconomic literature, the uncertainty is considered with regard to production decisions (Bloom, 2009; Bloom et al., 2018) or consumption and saving decisions (Fernandez-Villaverde et al., 2011; Kaplan and Violante, 2018). We add a spatial dimension and focus on decisions about jobs. Doing so changes the qualitative effects (from absolute to relative) as well as the quantitative effects of uncertainty, as the latter is a driver of comparative (dis-)advantage in the spatial economy.

The remainder of the paper is organized as follows. In the next section we model the intertemporal allocation of labor among jobs (i.e., region, sector and occupation) as a system of individual dynamic optimal control problems (DOCPs) augmented with a law of motion for the distribution of labor force. In the third section we describe the within-period equilibrium of the competitive multi-country multi-sector economy, given the pre-determined allocation of labor. In the fourth section we discuss the distinctive role that uncertainty plays in spatial models as opposed to the typical setup of macro models with heterogeneous agents, that is in closed economy. In the fifth section we discuss the methodology to obtain the numerical solution of the model. In section six we show how to take the model to the data and section seven outlines the results from a calibration and quantitative analysis of the cost of uncertainty based on French data. The last section concludes. The Appendix provides a rich set of background materials.

2 Inter-temporal allocation of labor

In this section we describe a dynamic general equilibrium model with spatially segmented markets, both for products and labor. We build on Caliendo et al. (2019) to characterize the spatial equilibrium and follow Artuç et al. (2010) to model how households of different type make discrete, forward-looking decisions on their job. We specifically allow for risk-averse households to face aggregate uncertainty about the value of jobs under rational expectations but imperfect foresight.

2.1 Endowments and market structure

We consider a world economy populated by a unitary measure of mobile households distributed in $r = 1, ..., R$ segmented regions. In each region, a large number of perfectly competitive firms produce goods that are horizontally differentiated by region of origin and sector $s = 1, ..., S$. Firms located in region r and producing in sector s combine a region-sector specific stock of $H^{rs} > 0$ units of local, immobile structures with mobile labor. The latter is located in the same region r, supplied in the same sector s, and differentiated by occupation type $k = 1, 2, ..., K$.

The region-sector-occupation triplet defines a job $j = \{r, s, k\}$. Every job is filled by a measure of agents who supply labor inelastically in exchange for the market clearing wage in the job-specific market. There are no means of saving, so that, in each period, households consume the equivalent of their real wage determined by their job-specific wage and their region-specific price index.

Time is discrete and denoted by $t = 0, 1, 2, ...$ for an infinite horizon. Within every period t, goods are traded at their marginal cost in region-segmented output markets, and jobs are allocated at the value of their marginal product in job-segmented labor markets. Between time periods, households reallocate optimally across jobs, as we will describe below. The measure of households in each job at the beginning of a period $t + 1$ is a predetermined outcome of the choice made by individuals before the end of period t about which job to choose in the next period.

2.2 Lifetime age spells

Every individual transits through four age spells during her lifetime, $a = \{b, y, o, d\}$. A newborn, $a = b$, is born with probability $\lambda^* \in (0, 1)$ to young individuals currently in a certain job. Newborns do not consume or supply labor but prepare themselves to become

next period with probability $\lambda^b = 1$ young individuals, $a = y$, who will supply labor in a certain job. A young household will become old, $a = o$, with probability $\lambda^y \in (0, 1)$ at the beginning of the next period but still work. Old individuals will die at probability $\lambda^o \in (0, 1)$.

Both young and old households supply labor in the job in which they start the period, they consume by purchasing goods in the same region, where their job is located (i.e., there is no commuting).[7] By the end of each period both young and old individuals make a decision on the job they will start with in the next period. However, upon dying old individuals drop out of the population at zero value. Therefore, older households have a shorter expected lifetime than young households.

2.3 Benefits and costs from reallocating across jobs

Households enjoy two sources of value: utility from consumption of goods in the current period t and satisfaction from the job that they will start with in the next period $t + 1$. Preferences across consumption bundles are the same for every household. Instead, individuals are heterogeneous in their individual satisfaction from having a certain job. Their objective is to maximize the discounted sum of utility from consumption and job satisfaction over their lifetime.

Heterogeneity about job satisfaction captures in a reduced form the non-pecuniary value of a given job across individuals. Every period households draw i.i.d. realizations of taste shocks for every job. They are distributed with c.d.f. $F(\varepsilon)$ and p.d.f. $f(\varepsilon)$. An individual h that finds herself in job j at the end of period t collects a value $\nu\varepsilon^j_{h,t}$, where $\nu > 0$ scales the contribution of idiosyncratic payoffs for being in a job j at the end of period t, $\varepsilon^j_{h,t}$.[8] The distribution of taste shocks is constant over time, individuals, and job characteristics such that $\varepsilon^j_{h,t} \sim F(\varepsilon)$.

Idiosyncratic tastes are contrasted with costs for changing the job that vary by the job of origin and the job of destination, and eventually even by age. Households of age $a = \{b, y, o\}$ face a cost $\zeta^{aj,j'} \geq 0$ of changing from the current job j to a future job j'. Moving costs are

[7] In the quantitative model we present below, this is justified by the relatively large size of regions we consider in France. The choice of this granularity is dictated by some of the data required for the model calibration. However, putting together regions, sectors, and skill levels, there will be more than $2,000$ jobs between which individuals can transit. Hence, the state space of the model will be large.

[8] An intuitive interpretation of ν is the degree of job differentiation: if jobs are perfect substitutes such that households are indifferent between them, i.e., there is no differentiation $\nu = 0$, then rational households would only be interested in reallocating across jobs that maximize their lifetime stream of consumption. Otherwise, an optimal reallocation across jobs is the outcome of balancing utility from consumption and job satisfaction over the lifetime. Thus, ν can be seen as capturing the intrinsic value of the option to move to a certain job conditional on observable characteristics.

measured in terms of utility loss, they are assumed to be time-invariant and non-stochastic.[9] This implies that the net benefit for a household h of moving from a job j at time t to a job i at time $t+1$ is given by $(\varepsilon^i_{h,t} - \zeta^{aj,i}) - (\varepsilon^j_{h,t} - \zeta^{aj,j})$, which consists of an idiosyncratic taste and a moving cost that is common to all households with the same age moving between the same jobs.

Frictions captured in the moving costs are exogenous to the model, but with respect to the dimension of age we impose a structure, namely $\zeta^{oj,j'} \geq \zeta^{yj,j'} \geq \zeta^{bj,j'}$ within the same job pair.[10] We choose this structure to ensure that the incentive for older individuals to stay in a less attractive job relative to young individuals is not systematically counteracted by an age-related decline in moving costs. Without loss of generality, we normalize the cost of remaining in the same job to be zero, $\zeta^{aj,j} = 0$ at every age $a = \{b, y, o\}$.

2.4 Inter-temporal problem of individuals

At the beginning of every period t, production starts given the current distribution of labor across jobs $L_t = \{L^j_t\}_{j \in J}$ and the aggregate stochastic state of the economy $Z_t = \{Z^j_t\}_{j \in J}$, that consists of realizations of aggregate shocks which affect the distribution of real income across jobs.[11] Through this channel, aggregate uncertainty affects the mobility between jobs.

Before the end of every period t a certain household h realizes the vector of idiosyncratic tastes for all jobs $\varepsilon_{h,t} = \{\varepsilon^j_{h,t}\}_{j \in J}$. Given the endogenous state L_t, the realization of the stochastic state Z_t and the realizations of job specific shocks $\varepsilon_{h,t}$, each individual makes an independent forward-looking decision about the optimal job where to end the period.

Let $W^{aj}(L_t, Z_t, \varepsilon_{h,t})$ be the lifetime welfare of a household of age a in job j with a vector of idiosyncratic values $\varepsilon_{h,t}$. Let $V^{aj}(L_t, Z_t) = \mathbb{E}_h[W^{aj}(L_t, Z_t, \varepsilon_{h,t})]$ be the average lifetime value of a household in age spell a and job j at time t, where the operator $\mathbb{E}_h[\cdot]$ yields the expectation taken over the idiosyncratic realizations of the value $\varepsilon^j_{h,t}$ across all households.

[9] As in frameworks that our model is related to (e.g., Dix-Carneiro, 2014 or Caliendo et al., 2019), moving costs summarize in a reduced form all frictions to changing jobs that the model is silent about.

[10] As said before, the newborns do not work or consume. We index the latter with an origin job j to indicate that they inherit an information set associated with their parents' job j that provides an initial condition in terms of where to start their employment journey, before choosing their first job. Given this information summarized in the moving costs $\zeta^{bj,j'}$, newborns can start in the same job as their parents or in another one.

[11] For concreteness, we anticipate that in the specific implementation of the model we will consider a stationary stochastic process of total factor productivity (TFP) at the level of region-sector pairs. However, the approach is not restricted to stationary technological shocks. The analysis goes through for any aggregate stochastic state with a Markov process that satisfies Feller's property, such that for every continuous, bounded and real-valued function of Z_{t+1} the expected value conditional on Z_t is continuous and bounded.

Let $\omega^{aj}(L_t, Z_t)$ be the utility from current consumption when the household of age a is in a job j and the aggregate state is $\{L_t, Z_t\}$. We assume that future values are discounted at a rate $\beta \in (0,1)$. The lifetime value of a rational household h conditional on its age and job status is given by:

$$W^{aj}(L_t, Z_t, \varepsilon_{h,t}) = \omega^{aj}(L_t, Z_t) + \max_{n \in J} \left\{ \nu \varepsilon^n_{h,t} - \zeta^{aj,n} + \beta \mathbb{E}_t \left[\bar{V}^{an}(L_{t+1}, Z_{t+1}) \right] \right\}, \quad (1)$$

where $\mathbb{E}_t[\cdot]$ is the expectation operator given the information available at time t. Assuming that the newborn do not earn nor consume and that the young and the old in the same job j earn the same wage and consume the same, we obtain

$$\omega^{bj}(L_t, Z_t) = 0, \quad \omega^{aj}(L_t, Z_t) = \omega^{j}(L_t, Z_t) \, \forall a \in \{y, o\},$$

where $\bar{V}^{aj}(L_{t+1}, Z_{t+1})$ denotes an average across future age spells and is defined as

$$\bar{V}^{aj}(L_{t+1}, Z_{t+1}) = (1 - \lambda^a) V^{aj}(L_{t+1}, Z_{t+1}) + \lambda^a V^{a+1,j}(L_{t+1}, Z_{t+1}) \, \forall a \in \{y, o\},$$

using $\bar{V}^{bj}(\cdot) = V^{yj}(\cdot)$, $V^{a+1,j}(\cdot) = V^{o,j}(\cdot)$ for $a = y$, and $V^{a+1,j}(\cdot) = 0$ for $a = o$.

We wish to characterize the average across households of values in (1) by age and job group, i.e., $V^{aj}(L_t, Z_t) = \mathbb{E}_h[W^{aj}(L_t, Z_t, \varepsilon_{h,t})]$. A convenient parametrization of the distribution of idiosyncratic tastes as Gumbel yields a closed-form solution for the discrete-choice problem in (1). A detailed derivation is discussed in Appendix A. The value of the DOCP for an individual in age spell a job j is given by

$$V^{aj}(L_t, Z_t) = \omega^{aj}(L_t, Z_t) + \nu \ln \left(\sum_{k \in J} e^{\frac{1}{\nu}\left(\beta \mathbb{E}_t\left[\bar{V}^{ai}(L_{t+1}, Z_{t+1})\right] - \zeta^{aj,i}\right)} \right) \quad (2)$$

and the associated optimal policy describing the fraction of movers to a job i is given by:

$$m_t^{aj,i} = \frac{e^{\frac{1}{\nu}\left(\beta \mathbb{E}_t\left[\bar{V}^{ai}(L_{t+1}, Z_{t+1})\right] - \zeta^{aj,i}\right)}}{\sum_{k \in J} e^{\frac{1}{\nu}\left(\beta \mathbb{E}_t\left[\bar{V}^{ak}(L_{t+1}, Z_{t+1})\right] - \zeta^{aj,k}\right)}}. \quad (3)$$

The solution in (2) generalizes the spatial dynamic problem of Caliendo et al. (2019) in two directions: households have rational expectations about an uncertain future; and they are heterogeneous with respect to an unavoidable risk associated with aging. Equation (3) determines the reallocation of households over jobs between time t (in which some households

leave their current job j) and the next period $t+1$ (in which these households are in their new job i). This information is what will be sufficient to characterize – in a deterministic way – the new allocation of households across jobs, L_{t+1}.

2.5 Law of motion of the labor force

Let $L_t^{bj} = \lambda^b L_t^{yj}$, L_t^{yj}, and L_t^{oj} be the populations of newborn, the young, and the old individuals in a job j at time t. With the introduced age-transition probabilities, these populations will turn into the following age groups in $t+1$. The young of $t+1$ without a job index, L_{t+1}^y will consist of $\lambda^b L_t^y$ "new young" and $(1-\lambda^y)L_t^y$ "remaining young". The old of $t+1$ without a job index, L_{t+1}^o will consist of $\lambda^y L_t^y$ "new old" and $(1-\lambda^o)L_t^o$ "remaining old" (not deceased). The individuals of these groups will have started out from some job i in t and transited to job j at $t+1$ at a rate $m_t^{ai,j}$ for age group a. Using the convenient notation of $\lambda^{y-1} = \lambda^b$, $\lambda^{o-1} = \lambda^y$, $m_t^{y-1i,j} = m_t^{bi,j}$, and $m_t^{o-1i,j} = m_t^{yi,j}$, the law of motion for individuals of age $a \in \{y, o\}$ can be written as:

$$L_{t+1}^{aj} = (1-\lambda^a)\sum_{i=1}^{J}\left(m_t^{ai,j}L_t^{ai}\right) + \lambda^{a-1}\sum_{i=1}^{J}\left(m_t^{a-1i,j}L_t^{a-1,i}\right). \tag{4}$$

We postulate the initial condition:[12]

$$L_0^o = \sum_{j=1}^{J} L_0^{oj} = \frac{\lambda^y}{\lambda^y + \lambda^o}. \tag{5}$$

Clearly, while this condition implies a stable aggregate population by design, the distribution of the young and the old across jobs (thus, across regions, sectors, and occupations) will be endogenous.

3 Within-period allocation of output

The DOCP of Section 2 at each period t takes the level of utility, $\omega^{aj}(L_t, Z_t)$ for each age $a = \{o, y\}$ and job $j = 1, ..., J$, that solves the within-period equilibrium of the model for a predetermined aggregate deterministic state L_t and a given realization of the aggregate stochastic state Z_t. We assume that the instantaneous utility features constant relative risk

[12] See Appendix A for a detailed discussion of the stationarity condition on population size.

aversion

$$\omega^{oj}(L_t, Z_t) = \omega^{yj}(L_t, Z_t) = \begin{cases} \ln(C_t^j / \min_i C_0^i) & \text{if } \chi = 1 \\ \frac{(C_t^j / \min_i C_0^i)^{1-\chi} - 1}{1-\chi} & \text{if } \chi \in (0,1) \end{cases} \tag{6}$$

where $C_t^j \equiv C^j(L_t, Z_t)$ is the aggregate consumption at time t by an agent in a job $j = \{r, s, k\}$ who supplies labor in region r, sector s, and occupation k, and $\chi \in (0, 1]$ is the coefficient of relative risk aversion, such that $1/\chi$ is the elasticity of intertemporal substitution.[13]

Agents allocate consumption across sectoral goods according to Cobb-Douglas preferences:

$$C_t^j = \prod_{s=1}^{S} \left(c_t^{j,s}\right)^{\alpha^s} \tag{7}$$

spending a share $\alpha^s > 0$ of their income, with $\sum_{s=1}^{S} \alpha^s = 1$, on the consumption of $c_t^{j,s} \equiv c^{j,s}(L_t, Z_t)$ units of the final good from sector s. Let w_t^j be the nominal income of an agent in a job $j = \{r, s, k\}$ and call P_t^{rs} the price index of final goods in region r sector s. Then, the consumption of final goods in region r and sector s' by an agent in job $j = \{r, s, k\}$ at time t is $c_t^{j,s'} = \alpha^{s'} w_t^j / P_t^{rs'}$.

Prices and consumption levels are determined in the within-period equilibrium given the allocation of labor and the realization of the aggregate stochastic state of the economy, thus, (L_t, Z_t) is the state of the economy at time t. Specifically, the market structure of the economy follows the one in Caliendo et al. (2018) and the aggregate stochastic state consists of region-sector-specific realizations $\{A_t^{rs}\}_{r=1,s=1}^{R,S}$ of total factor productivity (TFP). Hence, firms in a sector and region employ labor, structures, and intermediates to produce output which is used as an intermediate input in production or consumed as a final good.

The within-period equilibrium is characterized by a set of exogenous and time-invariant parameters: the cost shares for skill-specific labor in all labor costs (ϵ), the share of value added in all production costs (double-indexed γ), the share of structures in value added

[13] In the definition of utility (6), scaling for the minimum consumption in the initial period does not affect the optimal allocation, nor the sorting of utility across jobs over time, that matters for the intertemporal allocation. This normalization is convenient for tractability in a setup in which agents get old and eventually die, since it implies that the value of being matched at any job in the initial period is not lower than the value of dropping out of the population (zero, by construction).

(ξ), the share of specific intermediates in all intermediates (quadruple-indexed γ), trade imbalances (D), the endowment with structures (H), interregional and international trade costs (τ), the Fréchet parameter governing the productivity distribution (θ), consumption shares (α), and parameter composites (B and Γ). We outline this specific model in detail in Appendix B. We only report the key equilibrium conditions for the sake of brevity.

Given the exogenous parameters of the model and the realization of (L_t, Z_t), the within-period equilibrium consists of six sets of variables: wages $\{w_t^{rsk}\}_{r=1,s=1,k=1}^{R,S,K}$, rental prices for structures $\{\rho_t^{rs}\}_{r=1,s=1}^{R,S}$, sales $\{X_t^{rs}\}_{r=1,s=1}^{R,S}$, bilateral fractions of expenditure $\{\pi_t^{rs,r's}\}_{r=1,s=1,r'=1}^{R,S,R}$, local prices of aggregate goods $\{P_t^{rs}\}_{r=1,s=1}^{R,S}$ and relative local prices of intermediate goods $\{x_t^{rs}\}_{r=1,s=1}^{R,S}$ that solve the system of six equilibrium conditions:

$$
\begin{aligned}
w_t^{rsk} &= \frac{\epsilon_{sk}(1-\xi^{rs})\gamma^{rs}}{L_t^{rsk}} \sum_{r'=1}^{R} \pi_t^{r's,rs} X_t^{r's}, \qquad (8)\\
\rho_t^{rs} &= \frac{\xi^{rs}}{1-\xi^{rs}} \frac{L_t^{rs}}{H^{rs}} \bar{w}_t^{rs} \ , \text{ where } \bar{w}_t^{rs} = \left(\sum_k w_t^{rsk} L_t^{rsk} \right) / L_t^{rs} \text{ and } L_t^{rs} = \sum_k^K L_t^{rsk},\\
X_t^{rs} &= \sum_{s'=1}^{S} \gamma^{rs',rs} \left(\sum_{r'=1}^{R} \pi_t^{r's',rs'} X_t^{r's'} \right) + \alpha^s \left(\sum_{s'=1}^{S} \left(\bar{w}_t^{rs'} L_t^{rs'} + \rho_t^{rs'} H^{rs'} \right) + D_t^r \right),\\
\pi_t^{rs,r's} &= \frac{\left(x_t^{r's} \tau^{rs,r's} \right)^{-\theta^s} \left(A_t^{r's} \right)^{\theta^s}}{\sum_{r''=1}^{R} \left(x_t^{r''s} \tau^{rs,r''s} \right)^{-\theta^s} \left(A_t^{r''s} \right)^{\theta^s}},\\
P_t^{rs} &= \Gamma^{rs} \left(\sum_{r'=1}^{R} \left(x_t^{r's} \tau^{rs,r's} \right)^{-\theta^s} \left(A_t^{r's} \right)^{\theta^s} \right)^{-\frac{1}{\theta^s}},\\
x_t^{rs} &= B^{rs} \left[(\rho_t^{rs})^{\xi^{rs}} \left(\prod_{k=1}^{K} \left(w_t^{rsk} \right)^{\epsilon_{sk}} \right)^{1-\xi^{rs}} \right]^{\gamma^{rs}} \prod_{s'=1}^{S} \left(P_t^{rs'} \right)^{\gamma^{rs,rs'}}.
\end{aligned}
$$

4 Uncertainty

The characterization of the DOCP in Section 2, given a within-period equilibrium as the one described in Section 3, is completed by a random process for the aggregate stochastic state $Z_t = \{Z_t^j\}_{j \in J}$, that is anchored to the unpredicted (stochastic) component of an autoregressive process for region-sector-specific TFP levels. The one key feature of the latter is that it affects the perceived (certainty-equivalent) real income associated with a job, which is what agents compare when making their location decisions under uncertainty.

Inter-temporal optimality of the DOCP implies that for the marginal mover from job j

to job n the cost of moving equals the discounted expected total gains from moving, i.e., the difference in expected lifetime value of any job relative to the reference job an agent currently holds. The lifetime value consists of the current value and the option (or continuation) value, as can be seen from equation (2).

For an individual of a specific age, the expected lifetime value in a specific job inter alia depends on (i) shocks that she can arbitrage against by choosing optimally a job to start at in the next period, (ii) aging risk, which cannot be avoided and truncates the expected lifetime horizon as well as the continuation value, (iii) moving costs between jobs that are known at the time of reallocation. In this context, the role of uncertainty can be understood looking at four main channels.

1. *If agents are risk-averse, aggregate uncertainty depresses the incentive to relocate relative to perfect foresight.* Hence, the expected gains from moving under the adopted assumptions are lower for risk-averse agents with rational expectations under uncertainty.

2. *Due to aggregate uncertainty, more households rationally spend a greater portion of their life in relatively bad jobs.* In a population of rational but risk-averse agents, uncertainty reduces the incentive to relocate toward jobs that gain value in relative terms. Therefore, a greater fraction of households than under perfect foresight remains in deteriorating jobs.

3. *Aging reduces the option value of relocation.* As long as moving costs do not decline with age, the difference in expected lifetime implies that, ceteris paribus, the option value to reallocate to any job is greater for a young household than for an old one.

These three insights are familiar to macroeconomic studies that adopt an option-value approach to choices made by risk-averse agents under uncertainty, and they unambiguously qualify uncertainty as a welfare cost.[14] However, in a spatial equilibrium model with multiple jobs, these jobs are interdependent through intratemporal linkages (trade) as well as the intertemporal ones (mobility of agents across jobs). Then, uncertainty plays a broader role.

4. *Uncertainty as a source of comparative (dis-)advantage of jobs.* In a spatial open economy with multiple, interlinked jobs, what matters for the attractiveness of a job is ceteris paribus not the absolute degree of uncertainty associated with it but the one relative to other jobs. And in equilibrium, uncertainty interacts with other fundamentals determining the value of a job. Hence, the opportunity costs of moving to a certain job differ from what they would have been under perfect foresight.

[14] In the Appendix, Section D, we provide a more formal discussion of these results.

Through this channel, uncertainty may generate a source of comparative (dis)advantage across jobs. This creates room for new welfare gains in some jobs and welfare losses in others, due to how uncertainty shapes the distribution of labor and, in turn, of output production, across regions, sectors and occupations relative to a perfect-foresight world.

An illustration of the latter in a customary spatial equilibrium model is beyond the reach of analytical derivations. Therefore, we illustrate this mechanism through a numerical solution of the model, that is calibrated to France.

5 Solving the model

The system (2)-(4) describes a model in which heterogeneous agents decide about their job in the next period by solving a DOCP, given the current allocation of labor across jobs and the aggregate stochastic state of the economy, (L_t, Z_t). Each agent acts in a rational, non-cooperative way and takes into account that the next-period allocation of labor L_{t+1} results from the aggregation of current optimal choices. Such a dynamic model is well-suited to be characterized as a backward-forward Mean Field Game (MFG) of coupled Bellman equations (2), with policies (3), and Kolmogorov equations (4) under aggregate uncertainty.

Obtaining a solution for such a MFG with aggregate uncertainty is challenging. Only recently, the literature in mathematics has developed numerical approaches that provide a solution to MFGs in continuous time, continuous state, and in the absence of aggregate uncertainty. In their Section 3, Achdou et al. (2021) provide an important discussion of solution techniques of MFGs with a focus on economic problems. However, our model (i) is in discrete time, (ii) features discrete individual deterministic states (here, "jobs"), (iii) accounts for aggregate uncertainty, and (iv) keeps track of two overlapping generations of individuals who supply labor, $a \in \{o, y\}$.[15] Furthermore, the scale is large: the model consists of $R \cdot S$ region-sector pairs and K-many skills, so that there are $J = R \cdot S \cdot K$ jobs.

Tasks of the type at hand call for the development of a new solution algorithm, that we outline in the following subsections.

[15] The young are born and get old, and the old die, all at respective probabilities, so that the size of the population and the relative size of the age groups in the whole population remains constant, while the composition of the labor force by age across jobs can change over time.

5.1 Notation

Let $\mathcal{AJ} = (1, 2, ..., J, J+1, J+2, ..., 2 \cdot J)$ be the space of age and job pairs $\{(aj), \forall a = o, y, \forall j = 1, ..., J\}$. Denote as $\mathcal{F}^{2 \cdot J} = \{(L^1, L^2, ..., L^{aj}, ..., L^{2 \cdot J}) : L^{aj} \in (0,1) \forall (aj) \in \mathcal{AJ}, \sum_{(aj) \in \mathcal{AJ}} L^{aj} = 1\}$ the space of frequency distributions on $\mathcal{AJ}$. The aggregate deterministic state of the problem at time t is the vector $L_t = \{L_t^{aj}\}_{aj=1}^{2 \cdot J} \in \mathcal{F}^{2 \cdot J}$.

Define the compact set $\mathcal{Z}^{2 \cdot J} \subset \Re_+^{2 \cdot J}$ such that an element $Z \in \mathcal{Z}^{2 \cdot J}$ consists of a vector of $R \cdot S$ realizations of region and sector-specific TFP levels, each replicated $2 \cdot K$ times for age and occupation types employed in the same region-sector pair. The aggregate stochastic state of the problem at time t consists of a vector $Z_t = \{Z_t^{aj}\}_{aj=1}^{2 \cdot J} \in \mathcal{Z}^{2 \cdot J}$.

Define the random mapping $\mathcal{Q} : \mathcal{Z}^{2 \cdot J} \to \mathcal{Z}^{2 \cdot J}$. We assume that the random process for the aggregate stochastic state $Z_{t+1} = \mathcal{Q}(Z_t)$ for $t = 0, 1, ...$ is Markov and satisfies Feller's property.

5.2 Master Equation

Consider the Bellman equation (2). Given a non-empty and compact set of actions, a Lipschitz-continuous, bounded and positive-real valued utility function $\omega^{aj} : \mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J} \to \Re_+$, and an aggregate stochastic state with Markov and Feller properties, the value function $V^{aj} : \mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J} \to \Re_+$ is Lipschitz-continuous, bounded and positive real-valued. Call $\boldsymbol{V} = \{V^{aj} : \mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J} \to \Re_+, \forall (aj) \in \mathcal{AJ}\} \in \Re_+^{2 \cdot J}$ a $2 \cdot J$-dimensional vector of continuous, bounded and positive real-valued functions defined on $\mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J}$.

For a vector of value functions $\boldsymbol{V} \in \Re_+^{2 \cdot J}$, a frequency distribution $L_{t+1} \in \mathcal{F}^{2 \cdot J}$ and an aggregate stochastic state $Z_t \in \mathcal{Z}^{2 \cdot J}$, denote as $m_t^{aj,i} \equiv m^{aj,i}(\boldsymbol{V}, L_{t+1}, Z_t)$ the fraction of movers implied by the policy of the DOCP (3) given the state (L_t, Z_t). The system of Kolmogorov equations (4) defines a $2 \cdot J$-dimensional vector of continuous and bounded functions $\boldsymbol{L}_{next} = \{L_{next}^{aj} : \mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J} \to (0,1), \forall (aj) \in \mathcal{AJ}\} \in \mathcal{F}^{2 \cdot J}$ that satisfies the fixed-point $\boldsymbol{L}_{next}(L_t, Z_t) = M(\boldsymbol{V}, \boldsymbol{L}_{next}(L_t, Z_t), Z_t) L_t$ in $\mathcal{F}^{2 \cdot J}$, where $M(\boldsymbol{V}, L_{t+1}, Z_t)$ is a $2 \cdot J$-dimensional block-diagonal transition matrix, whose elements are, respectively, $\{(1 - \lambda^a) m^{ai,j}(\boldsymbol{V}, L_{t+1}, Z_t)\ (ai) \in \mathcal{AJ},\ j = 1, ..., J\}$ and $\{\lambda^a m^{a-1i,j}(\boldsymbol{V}, L_{t+1}, Z_t),\ (ai) \in \mathcal{AJ},\ j = 1, ..., J\}$.

Thus, a solution of the MFG consists of a pair of $2 \cdot J$-dimensional vectors of continuous and bounded functions $(\boldsymbol{V}^\star, \boldsymbol{L}_{next}^\star) \in \Re_+^{2 \cdot J} \times \mathcal{F}^{2 \cdot J}$ defined on the state space $\mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J}$ such that $\boldsymbol{L}_{next}^\star(L_t, Z_t) = M(\boldsymbol{V}^\star, \boldsymbol{L}_{next}^\star(L_t, Z_t), Z_t) L_t$ satisfies (2)-(4), while the economy is in equilibrium (6)-(8) at each state (L_t, Z_t) for every period $t = 0, 1, ...$, given an initial

allocation $(L_0, Z_0) \in \mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J}$ that satisfies (5), and given the stochastic process $Z_{t+1} = \mathcal{Q}(Z_t)$. This defines the *Master Equation*, that is the solution of the MFG defined on the product space of value functions and frequency-distribution functions $\Re_+^{2 \cdot J} \times \mathcal{F}^{2 \cdot J}$.

Before proceeding, note that the solution of the MFG can be written in terms of the payoffs of moving from a job, say j, to other jobs, say n, given the current age spell a contingent on the current state $\varphi_{aj,n}[L_t, Z_t] : \Re_+^{2 \cdot J} \times \mathcal{F}^{2 \cdot J} \to \Re$

$$\varphi_{aj,n}[L_t, Z_t](\boldsymbol{V}, \boldsymbol{L}_{next}) = \tag{9}$$
$$= \frac{1}{\nu} \left(\beta \mathbb{E}_t \left[(1 - \lambda^a) V^{an} \left(\boldsymbol{L}_{next}(L_t, Z_t), \mathcal{Q}(Z_t) \right) + \lambda^a V^{(a+1)n} \left(\boldsymbol{L}_{next}(L_t, Z_t), \mathcal{Q}(Z_t) \right) \right] - \zeta^{aj,n} \right).$$

The MFG (2)-(4) as a continuous and bounded function of contingent payoffs takes the form:

$$V^{aj}(L_t, Z_t) = \omega^{aj}(L_t, Z_t) + \nu \ln \left(\sum_{n=1}^{J} e^{\varphi_{aj,n}[L_t, Z_t](\boldsymbol{V}, \boldsymbol{L}_{next})} \right) \qquad \forall (aj) \in \mathcal{AJ} \tag{10}$$

$$L_{next}^{an}(L_t, Z_t) = (1 - \lambda^a) \sum_{j=1}^{J} \left(\frac{e^{\varphi_{aj,n}[L_t, Z_t](\boldsymbol{V}, \boldsymbol{L}_{next})}}{\sum_{i=1}^{J} e^{\varphi_{aj,i}[L_t, Z_t](\boldsymbol{V}, \boldsymbol{L}_{next})}} \right) L_t^{aj} + \tag{11}$$
$$+ \lambda^a \sum_{j=1}^{J} \left(\frac{e^{\varphi_{a-1j,n}[L_t, Z_t](\boldsymbol{V}, \boldsymbol{L}_{next})}}{\sum_{i=1}^{J} e^{\varphi_{a-1j,i}[L_t, Z_t](\boldsymbol{V}, \boldsymbol{L}_{next})}} \right) L_t^{a-1j} \qquad \forall (an) \in \mathcal{AJ}$$

$$\textit{for } L_{t+1} = \boldsymbol{L}_{next}(L_t, Z_t) \textit{ and } Z_{t+1} = \mathcal{Q}(Z_t) \textit{ given } (L_0, Z_0), \tag{12}$$

where (10) is the Bellman equation, (11) is the Kolmogorov equation, and (12) summarizes the recursive structure of the MFG and the given initial state.

5.3 Solution algorithm

The proposed solution algorithm constructs a finite-dimension approximation of the MFG (10)-(12) such that the solution can be computed as the unique fixed-point of a contraction mapping on a complete metric space defined on $\Re_+^{2 \cdot J} \times \mathcal{F}^{2 \cdot J}$, and rests on the following five pillars.

[1] Finite-dimension representation of a system of MFGs. Consider an integer index and number $p = 1, ..., P$ of distributions of agents across age and job pairs $l^p \in \mathcal{F}^{2 \cdot J}$. The set of elements $(l^1, l^2, ..., l^P)$ serves as a basis to discretize the support of the aggregate deterministic state in P-many points in each job. Analogously, consider an integer index and

number $q = 1, ..., Q$ of realizations of the aggregate stochastic state $z^q \in \mathcal{Z}^{2 \cdot J}$. The set of elements $(z^1, z^2, ..., z^Q)$ serves as a basis to discretize the support of the aggregate stochastic state in Q-many points in each location (or job).

This finite-dimensional representation is aimed at working as if there were a system of $P \cdot Q$-many MFGs, each starting at one grid-point of the aggregate state $(l^p, z^q) \in \mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J}$, for $p = 1, ..., P$ and $q = 1, ..., Q$.[16] Define $\boldsymbol{\mathcal{V}} = \{\mathcal{V}_{aj}^{pq} = V^{aj}(L_{next}^{aj}(l^p, z^q), \mathcal{Q}(z^q)), (aj) \in \mathcal{AJ} : p = 1, ..., P,\ q = 1, ..., Q\}$ and $\boldsymbol{\mathcal{L}} = \{\mathcal{L}_{aj}^{pq} = L_{next}^{aj}(l^p, z^q), (aj) \in \mathcal{AJ} : p = 1, ..., P,\ q = 1, ..., Q\}$ to be, respectively, the vector of value functions and the vector of next-period frequency distribution functions given the current-period states $\{(l^p, z^q) : p = 1, ..., P,\ q = 1, ..., Q\}$. Thus, $(\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{L}})$ is the unknown of the system of MFGs with initial states in the discretized state space.

[2] Piece-wise linear interpolation of contingent payoffs on the solution space. Let $\varphi_{aj,n}^{pq}(\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{L}})$ be the payoff (9) contingent on a point (l^p, z^q) of the state space, i.e., $\varphi_{aj,n}^{pq}(\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{L}}) \equiv \varphi_{aj,n}[l^p, z^q](\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{L}})$. Although one starts from points of the discretized state space with a guess for the solution $\{(\mathcal{V}^{pq}, l^p) : p = 1, ..., P,\ q = 1, ..., Q\}$, the next-period frequency distribution $\boldsymbol{\mathcal{L}}^{pq} = \boldsymbol{L}_{next}(l^p, z^q)$ might not belong to the set $(l^1, ..., l^P)$ and the next-period realization of the aggregate stochastic state $\mathcal{Q}(z^q)$ might not belong to the set $(z^1, ..., z^Q)$. Thus, also the next-period value function $\boldsymbol{V}(\boldsymbol{L}_{next}(l^p, z^q), \mathcal{Q}(z^q))$ might not belong to the set $\{\mathcal{V}^{pq} : p = 1, ..., P,\ q = 1, ..., Q\}$.

However, the guess can be used to interpolate the $P \cdot Q$ matrices of format $[2 \cdot J \times J]$ of contingent payoffs $\varphi_{aj,n}^{pq}(\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{L}}) : \Re_+^{2 \cdot J} \times \mathcal{F}^{2 \cdot J} \to \Re$ on the solution space $\Re_+^{2 \cdot J} \times \mathcal{F}^{2 \cdot J}$ with vertices in $\{(\mathcal{V}^{pq}, l^p) : p = 1, ..., P,\ q = 1, ..., Q\}$. To this purpose, define a P-dimensional vector of barycentric weights:

$$\{b_p(L) \in [0,1],\ b_p(l^p) = 1\ \forall p = 1, ..., P :\ \Sigma_{p=1}^{P} b_p(L) = 1\ , L \in \mathcal{F}^{2 \cdot J}\}$$

as follows

Definition. *Given a metric function $\delta : \mathcal{F}^{2 \cdot J} \times \mathcal{F}^{2 \cdot J} \to \Re_+$ the barycentric weights associated with the distribution $l^p \in (l^1, l^2, ..., l^P)$ as a function of an arbitrary frequency distribution*

[16] The only restriction that we impose, since we will interpolate the solution, is that the given initial allocation belongs to the grid, $l^1 = L_0$. The remaining $P - 1$ distributions in $(l^1, l^2, ..., l^P)$ and the Q vectors of realizations of the aggregate stochastic state $(z^1, z^2, ..., z^Q)$ can be arbitrarily chosen, as we will discuss later in the implementation.

$L \in \mathcal{F}^{2\cdot J}$ *are given in two steps: first, let* $(k_1, k_2) = \arg\min_{(k',k'')}\{\delta(l^{k'}, L) + \delta(l^{k''}, L)\}$*, then*

$$b_p(L) = \begin{cases} \frac{\delta(l^{k_2},L)}{\delta(l^{k_1},L)+\delta(l^{k_2},L)} & \text{if } p = k_1 \\ \frac{\delta(l^{k_1},L)}{\delta(l^{k_1},L)+\delta(l^{k_2},L)} & \text{if } p = k_2 \\ 0 & \text{otherwise.} \end{cases}$$

Thus, barycentric weights locate an arbitrary frequency distribution in the hyper-segment between the two closest distributions $\boldsymbol{l}_{1(p)}, \boldsymbol{l}_{2(p)} \in (l^1, l^2, ..., l^P)$ given the chosen metric function.[17]

With respect to the aggregate stochastic state, the expectation operator $\mathbb{E}_t$ can be described by means of a $Q \times Q$ matrix of time-invariant Markov probabilities: call $\pi_{qq'} \in (0, 1)$ the probability of the event $Z_{t+1} = z^{q'}$ conditional on $Z_t = z^q$. This is sufficient to define a piece-wise linear interpolation of the payoff function (9) contingent on a point (l^p, z^q) of the discretized state space for an arbitrary solution $(\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{L}}) \in \Re_+^{2\cdot J} \times \mathcal{F}^{2\cdot J}$

$$\varphi_{aj,n}^{pq}(\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{L}}) = \frac{1}{\nu}\left(\sum_{p'=1}^{P} b_{p'}(\boldsymbol{\mathcal{L}}^{pq}) \sum_{q'=1}^{Q} \pi_{qq'}\left(\beta\left[(1-\lambda^a)\,\mathcal{V}_{an}^{p'q'} + \lambda^a \mathcal{V}_{(a+1)n}^{p'q'}\right] - \zeta^{aj,n}\right)\right). \tag{13}$$

Grid-valued contingent payoffs (9) defined for an initial state that belongs to the discretized state space $(l^1, ..., l^P) \times (z^1, ..., z^Q)$ are approximated by the interpolation (13), for a given guess $\{(\mathcal{V}^{p'q'}, l^{p'}) : p' = 1, ..., P,\ q' = 1, ..., Q\}$.

[3] Numerical system of MFGs. For convenience of notation, consider the column stack of $P \cdot Q$ vectors of value functions in $\Re_+^{2\cdot J}$ and $P \cdot Q$ vectors of frequency distribution functions in $\mathcal{F}_+^{2\cdot J}$, each of dimension $N = 2 \cdot J \cdot P \cdot Q$, and use Ω to denote the corresponding compact subset in $\Re_+^{2\cdot N}$. Let the $2 \cdot N$-dimensional vector $\boldsymbol{X} = \{(\mathcal{V}^{pq}, l^p) : p = 1, ..., P,\ q = 1, ..., Q\} \in \Omega$ be a guess for the solution of the system of MFGs on the discretized state space. Then, we refer to the barycentric piece-wise linear interpolation of contingent payoffs (13) as $\varphi_{aj,n}^{pq}(\boldsymbol{X}) : \Omega \to \Re$.

The system of MFGs consists of $2 \cdot N$ equations in as many unknowns, which are the elements of the vector $\boldsymbol{X} \in \Omega$. The problem can be written by means of two continuous functions defined on the compact space Ω, namely $B_{aj}^{pq} : \Omega \to \Re$ for the grid-valued Bellman

[17] It shall be noted that weights are positive only for two distributions among $(l^1, l^2, ..., l^P)$ and that the metric function δ can be an arbitrary function, as long as it satisfies the properties of a distance, e.g., *sup* norm or the *earth mover's distance*.

equations and $K_{an}^{pq}:\Omega\rightarrow[0,1]$ for the grid-valued Kolmogorov equations:

$$B_{aj}^{pq}(\boldsymbol{X})=\omega^{aj}\left(l^{p},z^{q}\right)+\nu\ln\left(\sum_{n=1}^{J}e^{\varphi_{aj,n}^{pq}(\boldsymbol{X})}\right) \tag{14}$$

$$K_{an}^{pq}(\boldsymbol{X})=(1-\lambda^{a})\sum_{j=1}^{J}\left(\frac{e^{\varphi_{aj,n}^{pq}(\boldsymbol{X})}}{\sum_{m=1}^{J}e^{\varphi_{aj,m}^{pq}(\boldsymbol{X})}}\right)l_{aj}^{p}+\lambda^{(a-1)}\sum_{j=1}^{J}\left(\frac{e^{\varphi_{(a-1)j,n}^{pq}(\boldsymbol{X})}}{\sum_{m=1}^{J}e^{\varphi_{(a-1)j,m}^{pq}(\boldsymbol{X})}}\right)l_{(a-1)j}^{p} \tag{15}$$

$$\text{given }\left(l^{p},z^{q}\right)\ \forall p=1,...,P\ ,\ \ q=1,...,Q. \tag{16}$$

The system (14)-(16) is the numerical counterpart to the MFG (10)-(12) defined on the grid of $P\cdot Q$ points $(\mathcal{V}^{pq},l^{p})\in\Re_{+}^{2\cdot J}\times\mathcal{F}^{2\cdot J}$, each being a solution on the discretized state space $\{(l^{p},z^{q}):p=1,...,P,\ \ q=1,...,Q\}$.

[4] Contraction mapping. The fixed-point equations (14)-(15) define the operator $G:\Omega\rightarrow\Omega$ describing the system of MFGs with initial states in the discretized state space, whose image is the vector $G(\boldsymbol{X})=\{(B_{o1}^{11}(\boldsymbol{X}),...,B_{aj}^{PQ}(\boldsymbol{X}),...,B_{yJ}^{PQ}(\boldsymbol{X}),K_{o1}^{11}(\boldsymbol{X}),...,K_{aj}^{PQ}(\boldsymbol{X}),...,K_{yJ}^{PQ}(\boldsymbol{X}))\}$. The operator $G(\boldsymbol{X})$ is a contraction on Ω with respect to the sup norm; see Section C in the appendix for a detailed derivation of this result. Banach's Contraction-mapping Theorem implies that $G:\Omega\rightarrow\Omega$ has a unique fixed point $\boldsymbol{X}_{\star}\in\Omega$ that can be constructed by iteration starting from an arbitrary guess $\boldsymbol{X}\in\Omega$. The solution $\boldsymbol{X}_{\star}$ to the system of MFGs (14)-(15) yields a numerical value of the DOCP $\mathcal{V}_{aj}^{pq}=B_{aj}^{pq}(\boldsymbol{X}_{\star})\in\Re_{+}$ and a numerical next-period frequency distribution $\mathcal{L}_{aj}^{pq}=K_{aj}^{pq}(\boldsymbol{X}_{\star})\in(0,1)$ for all $(aj)\in\mathcal{AJ}$, $p=1,...,P$ and $q=1,...,Q$, corresponding to the contingent payoffs $\varphi_{aj,n}^{pq}(\boldsymbol{X}_{\star})$ in (13).

The solution of the MFG (14)-(15) for a given initial state (L_0,Z_0) prescribes the deterministic next-period distribution L_1. Hence, steps [1]-[4] of the algorithm can be repeated for a next initial allocation (L_1,Z_1), and so on. Thus, the numerical solution of the MFG for an arbitrary number of periods ahead can be computed recursively. However, the algorithm delivers a solution not only to one MFG but to a system of $P\cdot Q$ MFGs, each characterized by a different initial allocation. This is a computational advantage since it allows to approximate the solution for an arbitrary initial allocation by interpolating across solutions of the different MFGs.

[5] Piece-wise linear interpolation of the MFG on the state space. For an arbitrary initial state $(L_t, z^q) \in \mathcal{F}^{2 \cdot J} \times \mathcal{Z}^{2 \cdot J}$, contingent payoffs (9) are interpolated on the discretized state space, with vertices in (13). The barycentric piece-wise linear function interpolating contingent payoffs on the support of the aggregate deterministic state is given by:

$$pwl[\varphi_{aj,n}]\left(L_t, z^q\right) = \sum_{p=1}^{P} b_p(L_t) \varphi_{aj,n}^{pq}(\boldsymbol{X}_\star), \tag{17}$$

which approximates the contingent payoffs arbitrarily well as the number of grid-points for the aggregate deterministic state P grows. Piece-wise linear functions of contingent payoffs (17) are then used to approximate the value function and the law of motion of the MFG with initial state (L_t, z^q)

$$num[V^{aj}]\left(L_t, z^q\right) = \omega^{aj}\left(L_t, z^q\right) + \nu \ln\left(\sum_{n=1}^{J} e^{pwl[\varphi_{aj,n}](L_t, z^q)}\right) \tag{18}$$

$$num[L_{next}^{aj}]\left(L_t, z^q\right) = \tag{19}$$

$$= (1-\lambda^a) \sum_{j=1}^{J} \left(\frac{e^{pwl[\varphi_{aj,n}](L_t, z^q)}}{\sum_{m=1}^{J} e^{pwl[\varphi_{aj,m}](L_t, z^q)}}\right) l_{aj}^p + \lambda^{(a-1)} \sum_{j=1}^{J} \left(\frac{e^{pwl[\varphi_{(a-1)j,n}](L_t, z^q)}}{\sum_{m=1}^{J} e^{pwl[\varphi_{(a-1)j,m}](L_t, z^q)}}\right) l_{(a-1)j}^p$$

$$given\ L_t,\ (l^p, z^q)\ for\ p = 1, ..., P\ ,\ q = 1, ..., Q, \tag{20}$$

which yields a numerical MFG that preserves the non-linearity of the exact MFG (10)-(12). The accuracy of the approximated solution can be checked substituting for the vectors of numerical solutions $num[\boldsymbol{V}]\left(L_t, z^q\right)$ and $num[\boldsymbol{L}_{next}]\left(L_t, z^q\right)$ in (10)-(11). This yields the benchmark functions for the Bellman equations

$$\check{V}^{aj}\left(L_t, z^q\right) = \tag{21}$$

$$\omega^{aj}\left(L_t, z^q\right) + \nu \ln\left(\sum_{n=1}^{J} e^{\frac{1}{\nu}\left(\beta \sum_{q'=1}^{Q} \pi_{qq'}\left[(1-\lambda^a) num[V^{an}]\left(num[\boldsymbol{L}_{next}](L_t, z^q), z^{q'}\right) + \lambda^a num[V^{(a+1)n}]\left(num[\boldsymbol{L}_{next}](L_t, z^q), z^{q'}\right)\right] - \zeta^{aj,n}\right)}\right)$$

and for the Kolmogorov equations

$$\check{L}_{next}^{an}\left(L_t, z^q\right) = \tag{22}$$
$$(1-\lambda^a)\sum_{j=1}^{J}\left(\frac{e^{\frac{1}{\nu}\left(\beta\sum_{q'=1}^{Q}\pi_{qq'}\left[(1-\lambda^a)num[V^{an}]\left(pwl[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)+\lambda^a num[V^{(a+1)n}]\left(num[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)\right]-\zeta^{aj,n}\right)}}{\sum_{i=1}^{J}e^{\frac{1}{\nu}\left(\beta\sum_{q'=1}^{Q}\pi_{qq'}\left[(1-\lambda^a)num[V^{ai}]\left(num[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)+\lambda^a num[V^{(a+1)i}]\left(num[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)\right]-\zeta^{aj,i}\right)}}\right)L_t^{aj}+$$
$$+\lambda^a\sum_{j=1}^{J}\left(\frac{e^{\frac{1}{\nu}\left(\beta\sum_{q'=1}^{Q}\pi_{qq'}\left[\left(1-\lambda^{a-1}\right)num[V^{a-1,n}]\left(num[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)+\lambda^{a-1}num[V^{an}]\left(num[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)\right]-\zeta^{a-1,j,n}\right)}}{\sum_{i=1}^{J}e^{\frac{1}{\nu}\left(\beta\sum_{q'=1}^{Q}\pi_{qq'}\left[(1-\lambda^{a-1})num[V^{a-1,i}]\left(num[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)+\lambda^{a-1}num[V^{ai}]\left(num[\boldsymbol{L}_{next}](L_t,z^q),z^{q'}\right)\right]-\zeta^{aj,i}\right)}}\right)L_t^{a-1j}$$

that allow the respective approximation errors to be computed:

$$\xi_V\left(L_t, z^q\right) = \left\|num[\boldsymbol{V}]\left(L_t, z^q\right) - \check{\boldsymbol{V}}\left(L_t, z^q\right)\right\|, \tag{23}$$
$$\xi_L\left(L_t, z^q\right) = \left\|num[\boldsymbol{L}_{next}]\left(L_t, z^q\right) - \check{\boldsymbol{L}}_{next}\left(L_t, z^q\right)\right\|. \tag{24}$$

If the approximation errors are below a pre-specified level of tolerance, we have reached a solution at the required precision. Otherwise, we increase the number of distributions P. A finer grid of distributions increases the accuracy at which the piece-wise linear interpolation of the system of MFGs (18)-(20) approximates the solution to the MFG (10)-(12) outside the grid points $(l^1, ..., l^P)$.[18]

5.4 Comparison with existing approaches

Before taking the model to the data and presenting the quantification results, we pause to compare our approach to the solution of dynamic spatial heterogeneous agents models with other very recent developments. We discuss five alternative specifications, by addressing the corresponding seminal contributions. We briefly comment on similarities and differences with our methodology; we refer the reader to the cited work for a more complete assessment of the methodologies.

a) Eigenvalue decomposition. Kleinman et al. (2021) develop a framework for the analysis of a dynamic discrete-choice migration model in which agents make forward-looking investment decisions in discrete time. The approach is based on the following steps: (i) characterization of the deterministic steady state; (ii) linearization around the steady state; (iii) eigenvalue decomposition to characterize stationary transitional dynamics. This setup is the closest (among those that we consider here) to the macroeconomic literature on heterogeneous agents models. The great advantage is tractability (both analytically and in the

[18] We discuss more about the intuition later, when we construct a quantification exercise that is parsimonious enough to illustrate how the algorithm works; see paragraphs 7.2 and 7.3.

interpretation of shocks), but this comes at the cost of considering only first-order changes in the economic fundamentals, while second-order nonlinear terms drop out of the linearized solution. However, the latter is at the heart of the present paper's interest. Thus, to answer our research question we must preserve the full non-linear structure of the model.

b) Sequence-space Jacobian. Auclert et al. (2021) provide an approach to solve general-equilibrium heterogeneous-agents models that develop further the perturbation method introduced by Reiter (2009). The latter approach is based on a linear perturbation of the deterministic steady-state solution of a model with aggregate shocks. The innovation of Auclert et al. (2021) is to write the problem not in a state-space representation (as in Reiter, 2009) but in terms of the full set of derivatives of equilibrium mappings around the steady state. The obtained sequence of Jacobians is a sufficient statistic to compute deviations from the steady state and the structure of the solution, which is linear in aggregates, guarantees a fast solution algorithm.

This design is well suited for a typical macroeconomic context, in which the aggregate outcome of individual decisions is summarized in a small set of prices. The key element of the approach, a Jacobian, is appropriate to work with when the set of actions can be summarized by continuous variables and their marginal increments. It is less suited for a large spatial dynamic model, thus, for being used with a reallocation over a set of categorical actions. Furthermore, this approach, as the previous one, relies on the existence of a steady state around which the model is linearized. Thus, it is not tailored to account for the analysis of uncertainty in highly non-linear models.

c) Continuous-time MFG. Bilal (2023) proposes a state-space approach that represents the economy as a continuous-time MFG and treats aggregate uncertainty using a perturbation method. First, the distribution of agents among heterogeneous allocations is treated as an explicit state variable, thus, the MFG is a set of value functions defined on the space of distributions and on the space of the aggregate stochastic state. Second, the dependence on the aggregate stochastic state is simplified by considering perturbations around a deterministic steady state (higher-order perturbations are feasible, to serve the purpose of dealing with aggregate uncertainty). Third, the Master Equation in the form of analytical local perturbations around the deterministic steady state is approximated by a finite-dimension representation on the discretized state space.

Our approach is also based on a state-space representation of a MFG defined on the space of distributions and on its finite-dimension representation. However, the described methodology requires in addition to our setup: existence of a steady state; a differentiable value

function describing the DOCP (in addition to be continuous, as we have it); infinitesimally small time interval between making a decision and the aggregate consequences of that decision. The last assumption entails the largest difference to our approach. Our goal is to solve a discrete-time MFG, whereas the methodology presented in Bilal (2023) is rooted in the mathematics literature on continuous-time MFGs.[19] Hence, this approach is not applicable to our research question. Computational advantages in solving MFGs in continuous time have been already documented in Achdou et al. (2021). But, in our context, their use would come at the cost of restricting significantly the nature of uncertainty we would be able to address.

d) Perturbation around the transition path. Fan et al. (2023) propose a local approximation of a spatial dynamic stochastic equilibrium around the transition path predicted under perfect foresight. In contrast to other perturbation methods, it does not rely on the existence of a steady state, as is the case with our approach. Two further similarities are the characterization in discrete time and the tractability of aggregate uncertainty, there by means of higher-order perturbations.

The fundamental difference to our approach is the representation of the equilibrium in the sequence space as a local perturbation around the perfect-foresight path, while our setup defines the economy as a MFG on the state space. This implies, first, that the law of motion of the distribution is not a fixed point (that is the "verification role" played by the Kolmogorov equation in our setup); and, second, value functions and policy functions there are expressed as Taylor expansions of deviations along the stochastic path from the deterministic path, while in our setup they are not approximated, neither around the steady state nor around the deterministic transition path. Different goals motivate the two approaches: their setup is designed for an equilibrium in which agents have believes about future exogenous states; our approach characterizes the evolution of the economy for agents who take into account the simultaneous deterministic evolution of the spatial distribution while making their decisions.

e) Machine learning. Han et al. (2022) propose an algorithm for the numerical solution of large-scale heterogeneous agents models with aggregate shocks based on deep learning. In the spirit of Krusell and Smith (1998), value functions and policies are defined on a set of moments of the distribution of agents over the state (and not on the space of the distribution,

[19] For the sake of intuition, as solving a partial differential equation is a different task from solving a difference equation, solutions to discrete-time MFGs are based on a different (and much scarcer) set of results than their continuous-time analogues (for instance, the definition of a weak solution to the associated Hamilton-Jacobi-Bellman equation – a partial differential equation – does not apply in discrete time).

as in our case). Given this approximation, value functions and policy functions are computed as solutions to the DOCP on simulated paths, in which moments of the distribution are computed by means of neural networks.

The adoption of machine learning makes the approach efficient and flexible, although this assessment is more robust when the state space is continuous (as in the canonical macro setup that the machine-learning approach is designed for) rather than with a categorical state space (as in spatial models). But this computational advantage comes at the cost of solving a simpler problem than the one at hand, in which moments of the distribution become sufficient statistics, ultimately selected by the neural network.

In essence, the approach that we propose solves a system of discrete-time MFGs by means of an explicit scheme on the coupled system of Bellman equations and Kolmogorov equations, defined on the finite-dimension representation of the state space. In comparison to a canonical perturbation method, the local approximation around a steady state (or a deterministic path) is replaced by a barycentric interpolation on a grid of initial allocations, while the discrete nature of the state allows for solving the Bellman equation, thus preserving the non-linearity of the problem. Barycentric interpolation and functional form of the Bellman equation are also sufficient to characterize the system of numerical MFGs as a contraction mapping on a complete metric space. Then, an exact solution to the grid-valued system of MFGs (14)-(16) is obtained by fixed-point iteration.

6 Taking the model to the data

We calibrate and simulate the model based on data of a specific country, France, which we consider to be open to trade with the Rest of the World (ROW). Specifically, we consider $R = 22$ French regions at the NUTS2 level in Eurostat's nomenclature plus one constructed ROW, $S = 49$ productive sectors corresponding to the classification used in WIOD, and $K = 2$ occupational types. Regarding labor types, we distinguish between high-skilled labor (defined as managers and professionals) and low-skilled labor (the rest). Data are collected at an annual frequency between 2003 and 2014.[20] Adding the skill (or worker-type) dimension

[20] We consider French administrative *régions* located in continental Europe, thus, excluding *Corse* and *Départements d'Outre Mer* (the overseas departments). Among the WIOD sectors, we exclude primary activities, mining, repairing services, arts and households activities. The information on characteristics of the French population of workers and firms comes from two administrative datasets: *Déclaration annuelle de données sociales (DADS)* and *Statistique structurelle annuelle d'entreprises*, (Ficus-Fare). The classification of labor types is based on the *Nomenclatures des professions et catégories socioprofessionnelles*, published by the French statistical office INSEE, in which managers and professionals correspond to skill groups 3 and

k leads to $2,254$ jobs in total, of which 98 are located in the ROW. While workers throughout the world matter for trade flows and, generally, supply and demand of outputs and other inputs, we focus on the mobility of labor among the $J = 2,156$ jobs within France.

In our data, we consider workers with age between 18 and 67 years in 2012. Among those, we distinguish the two aforementioned age groups of workers, the "*young*" defined as to be of an age below 40 years and the "*old*" with an age higher than 40 years. In our baseline year, 40.8% of the French workforce is young and 59.2% is old. Accordingly, we fix the idiosyncratic Poisson rates that discipline aging in the model such that the average duration of an age spell for being young is 20 years. I.e., transiting to an old age involves an arrival rate of $\lambda^y = 0.05$. The average duration of an age spell for being old is 29 years, with old agents dying and dropping out of the population with an arrival rate of $\lambda^o = 0.0345$. Given this parameterization, the rate at which the young cohort generates newborns is fixed at an arrival rate of $\lambda^* = 0.05$ such that the total population remains constant given the observed composition by age in 2012.

Table 1: Arrival rates of idiosyncratic age shocks

parameter	description	value	motivation
λ^*	arr. rate newborns	5%	constant aggregate population
λ^b	arr. rate newborns turning young	100%	constant aggregate population
λ^y	arr. rate aging for young	5%	avg. duration young age spell is 20 years
λ^o	arr. rate aging for old	3.45%	avg. duration old age spell is 29 years

We describe the construction of all data and the procedures to obtain parameters in detail in Appendix E. In the following, we give an intuition for the data employed, the identifying variation exploited, and we present summary statistics on the parameters and fundamentals used in the quantification exercise.

The quantification of the model aims at being in line with key economic fundamentals in France in 2012. Based on all estimated fundamentals, the calibration of trade deficits will ensure that real wages obtained in a within-period equilibrium given the actual distribution of the labor force in France in 2012 and fundamentals match real wages observed in France in that year. Note that we do not assume that this allocation of labor is a steady state. It is a strength of our approach that we do not have to enforce this. Importantly, the aim of our quantitative exercise is to assess the role of aggregate uncertainty. For this purpose, we compare the evolution of the distribution of labor and lifetime welfare across jobs between

4 of the ISCO-08 classification published by the International Labour Organization.

two alternative scenarios: perfect foresight versus uncertainty, while keeping the levels of fundamentals (cost-share parameters, TFP levels, endowments, trade costs, and moving costs) constant throughout.

6.1 Data requirements to quantify the model

The data sets employed to obtain the parameters governing the within-period equilibrium include the annual structural statistics of companies for France (the *FICUS-FARE* dataset published by INSEE), the French administrative employer-employee dataset *Déclaration Annuelle des Données Sociales* published by INSEE, and the WIOD data on input-output tables as well as their data on socioeconomic accounts and Eurostat's European Road Freight Transport Survey. Apart from these data, we need data on bilateral sales (trade flows) between all pairs of NUTS2 regions in France and the Rest of the World (ROW) ($R = 22 + 1$) for every sector (WIOD sector $S = 49$). Moreover, we use data from Eurostat to measure structures by region and sector. We will say more on the latter below.

For the intertemporal equilibrium, we also rely on the French administrative employer-employee dataset *Déclaration Annuelle des Donnees Sociales*, as this dataset provides information on the employees and their wage in a given job j (i.e., region r, sector s, and occupation/skill level k) in a given year t. Moreover, it contains information on the same employees on their job j in the previous year $t-1$ as well as the wage earned then. Using data for the pair of consecutive years of 2009-2010 and 2012-2013, we can compute transitions between every pair of jobs j and n as well as associated wages for about 27 mln. workers per year.

6.2 Observed spatial asymmetries

Beyond few parameters that are common across French regions and only vary at the sectoral level (trade elasticities, θ^s, and consumption shares, α^s), most parameters of the model lead to spatial asymmetries across French regions. The model features five channels of spatial asymmetries: (i) time-invariant factor-cost share parameters; (ii) frictions to the mobility of agents, captured by time-invariant moving costs between all $j = \{rsk\}$- and $n = \{r's'k'\}$-indexed job pairs, ζ^{jn}; (iii) trade frictions within France and internationally, captured by time-invariant trade costs between all rs-indexed region-sector pairs, $\tau^{rs,r's}$; (iv) region-sector-specific endowments, captured by time-invariant stocks of structures H^{rs}, and the initial distribution of labor by type across regions and sector, L_0^{rsk}; (v) region-sector-

specific, time-varying total factor productivity (TFP), A_t^{rs}. Agents are uncertain about future realizations of TFP. Therefore, after calibrating overall TFP, we isolate its stochastic component, a_t^{rs}, from the deterministic one, and we model its dynamic stochastic process at the region-sector level.

The next five numbered paragraphs are dedicated one each to these ingredients. Together with consumption and trade-elasticity parameters, these ingredients inform exhaustively the given fundamentals of the model: agents are assumed to be rational, thus, they "know the model", including frictions, endowments, the deterministic technology component, and the stochastic process of technology shocks.

***(i)* Consumption shares, trade elasticity, and factor-cost shares.** In order to inform the cost-share parameters of the model (ϵ^{sk}, ξ^{rs}, γ^{rs}, and $\gamma^{rs,rs'}$), we rely on worker, accounting, and input-output data for France. Table 2 describes values of parameters for production and consumption that vary at the level of 49 sectors or at the level of 1,127 region-sector pairs, of which 1,078 are within France. The Table contains information on consumption shares and the trade elasticity that are obtained from WIOD data and taken from Caliendo and Parro (2015), respectively.[21]

Table 2: Summary statistics for consumption and production parameters

	# units	mean	std.dev.	p10	p50	p90
Consumption share α^s	49	0.020	0.035	0.001	0.006	0.071
Trade elasticity θ^s	49	5.816	6.869	2.550	4.550	9.270
Labor share $\gamma^{rs}(1-\xi^{rs})$	1,127	0.267	0.132	0.113	0.245	0.452
Structures share $\gamma^{rs}\xi^{rs}$	1,127	0.425	0.169	0.225	0.398	0.689
Materials share $\sum_{s'=1}^{S}\gamma^{rs,rs'}$	1,127	0.307	0.215	0.046	0.297	0.602

Note: The statistics "# units", "mean", "std.dev.", "p10", "p50", and "p90" refer to the number of units at which the data vary, the mean, the standard deviation, and the 10th, 50th, and 90th percentiles of the data across the units.

***(ii)* Moving costs.** Moving costs are estimated from observed job-to-job transitions in France, following the 2-stage procedure proposed in Artuç and McLaren (2015).[22] First,

[21] Details on the measurement of all parameters are provided in Appendix E.

[22] For a detailed discussion of the methodology, we refer the reader to Artuç and McLaren (2015). In Appendix E we report more details about our application of it.

Equation (3) capturing the optimal policy in terms of moving shares is estimated by Poisson pseudo-maximum-likelihood (PPML) estimation: origin-year and destination-year fixed effects capture the expected continuation value of jobs, moving costs are parameterized and fitted using geographical distance between regions and some binary indicators for switching region, sector, or occupation. In a first stage, this yields an estimate of moving costs, which is scaled by β/ν. In a second stage, an estimate for the inverse of the moving elasticity $\nu = 4.5281$ is obtained by fixing the discount factor parameter to $\beta = 0.95$ and using the model-implied utility based on measured real wages and an assumed form of utility (here, log-utility).

Table 3: Summary statistics for moving costs

	mean	std.dev.	p10	p50	p90
Mov. costs overall $\zeta^{rsk,r's'k'}$	118.13	56.58	22.46	150.57	154.07
Mov. costs between regions $\zeta^{rsk,r'sk}$	12.10	0.80	11.06	12.23	12.94
Mov. costs between sectors $\zeta^{rsk,rs'k}$	110.40	59.29	10.13	144.48	147.80
Mov. costs between occupations $\zeta^{rsk,rsk'}$	8.93	0.19	8.73	8.92	9.18

Note: The reported estimates refer to a log-utility specification with $\nu = 4.5281$ and $\beta = 0.95$. The statistics "mean", "std.dev.", "p10", "p50", and "p90" refer to the average, the standard deviation, and the 10th, 50th, and 90th percentiles of the data across the units.

Table 3 provides descriptive statistics of the estimated moving costs. Job-to-job flows between 2012 and 2013 indicate that, on average, 98.2% of the agents in a given job do not move to other jobs from one year to another. Overall, moving costs are smaller for changing occupation only within the same region and sector. In the subsample with moving costs below the 10th percentile cutoff, moving costs are higher for changing the sector, and they are slightly higher for changing the region. In the rest of the distribution, mobility frictions increase when changing the sector, while the cost of moving between regions or occupations remains flat.

***(iii)* Trade costs.** We calculate trade costs by means of the Head-Ries index (Head and Ries, 2001) using interregional sales and purchases between regions r and r' in sector s, $X^{rs,r's}$ and $X^{r's,rs}$, normalized by the intraregional absorption, $X^{rs,rs}$ and $X^{r's,'s}$:

$$\tau^{rs,r's} = \left(\frac{X^{rs,rs} X^{r's,r's}}{X^{rs,r's} X^{r's,rs}} \right)^{1/(2\theta^s)}.$$

Computations are conducted for the year 2012, and interregional flows between regions within France are based on Eurostat's European Road Freight Transport Survey (ERFTS).

Table 4: Summary statistics for trade costs

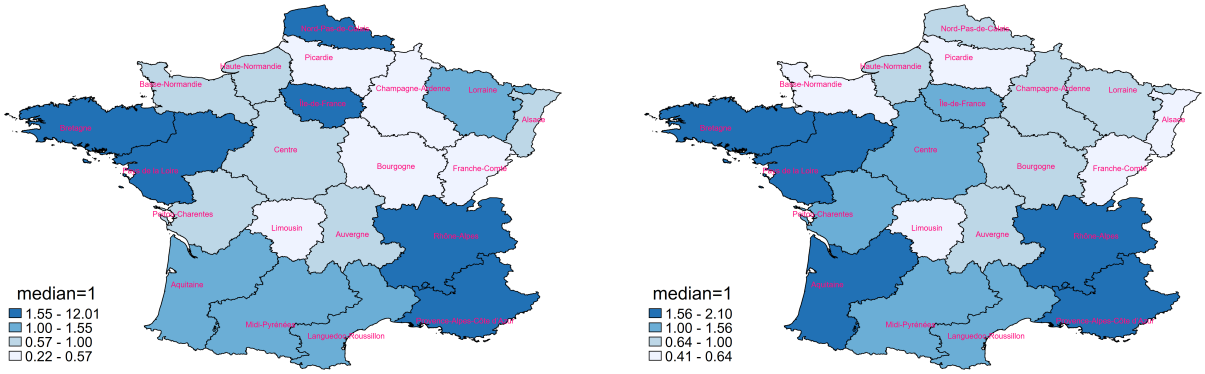
	mean	std.dev.	p10	p50	p90	V_{reg}	V_{sec}
Trade costs within France $\tau^{rs,r's}$	2.913	2.498	1.251	2.282	4.899	46.9%	26.8%
Trade costs with ROW $\tau^{rs,ROWs}$	25.981	134.869	2.763	5.535	15.452	7.5%	58.7%

Note: Own computations based on WIOD, Eurostat's ERFTS and Caliendo and Parro (2015). The statistics "mean", "std.dev.", "p10", "p50", and "p90" refer to the average, the standard deviation, and the 10th, 50th, and 90th percentiles of the data across the units. V_{reg} and V_{sec} are the percentages of the variation in trade costs explained by region- and sector-fixed effects, respectively.

Table 4 provides descriptive statistics on computed trade costs. Barriers within France are not negligible, although the costs of trading with the ROW are at least twice as high for any French region and on average more than eight times higher than for domestic interregional trade. The last two columns report on a variance decomposition, showing the percentage of the variance in trade costs explained by regions, V_{reg}, and sectors, V_{sec}. Within France, the regional component explains most of the variation, while trade costs of French regions and sectors with the ROW are mostly explained by the sectoral variation.

***(iv)* Endowments.** From French administrative matched employer-employee data (DADS) and balance-sheet data (Ficus-Fare) we compute the fraction of the employment of labor in hours by region, sector, and occupation in the baseline year, that is L^{rsk}_{2012}. Panel (a) in Figure 6.1 reports on the distribution of the labor bundle $L^{rs}_t \equiv (L^{rs1}_{2012})^{\epsilon_{sk}}(L^{rs0}_{2012})^{1-\epsilon_{sk}}$ across the French regions, where we weight each sector's contribution by the employment share of this sector in a region. Using two sources of data in Eurostat, *Land Use Overview by NUTS 2 Regions* and *Annual National Accounts, Breakdowns of Non-financial Assets by Type, Industry and Sector*, we measure the land use in square kilometers by sector and region in France. Using data on total fixed assets (in gross terms, according to the European System of Accounts 2010) by sector in the country, we use the share of assets by sector times sectoral and regional land use in the year 2012 as our measure for the units of structures used by region and sector, H^{rs}. In Panel (b) of Figure 6.1 we report on the distribution of structures across French regions averaged over sectors using sector-level employment weights.

***v*) Productivity.** In measuring region-sector-specific TFP, we follow Caliendo et al. (2018), whose intratemporal production equilibrium is fully consistent with the one considered



(a) Labor in 2012 (b) Structures in 2012

Figure 6.1: Sector-weighted average distribution of production factors across French regions

here.[23] Based on the aggregate sector-region-level production function implied by the within-period equilibrium, we calculate the region-sector-specific TFP as a Solow residual, using data on aggregate output, Y_t^{rs}, inputs, $H_t^{rs}, L_t^{rs}, M_t^{rs}$, price indices, P_t^{rs}, domestic absorption, $\pi_t^{rs,rs}$, as well as production function parameters to identify all components in real terms:[24]

$$\frac{Y_t^{rs}}{P_t^{rs}} = \frac{A_t^{rs}}{\Gamma^{rs}\left(\pi_t^{rs,rs}\right)^{\frac{1}{\theta^s}}}\left[\left(H_t^{rs}\right)^{\gamma^{rs}\xi^{rs}}\left(L_t^{rs}\right)^{\gamma^{rs}(1-\xi^{rs})}\left(M_t^{rs}\right)^{1-\gamma^{rs}}\right]. \tag{25}$$

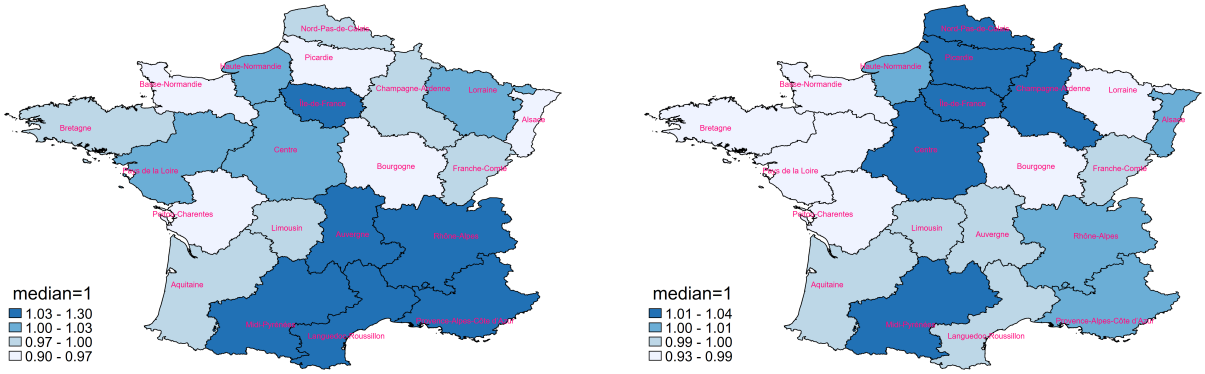
Panel (a) of Figure 6.2 reports on the distribution of measured TFP, A_t^{rs}, across French regions aggregated over sectors using employment-based weights. According to the figure, the south of the country (where manufacturing dominates) and Île de France, the region of Paris (the center of financial services), emerge as the areas with the highest productivity levels.

The proposed model is flexible regarding the origins of TFP. We specify its level as a function of time-invariant components that vary at the sector and region level, annual variations that are common across France, and a component that is a function of labor employed in a region and sector (see, e.g., Desmet et al., 2018 for the latter). This specification treats TFP as a function of the pre-determined agglomeration of labor L_t^{rs} in a region-sector.[25] To

[23] In particular, we define trade flows and value-added shares exactly as in Sections 3.3 and 4 in Caliendo and Parro (2015), and we follow Section 4 in Caliendo et al. (2018) in measuring TFP.

[24] All details on the derivation of Equation (25) and the data employed are presented in Appendix E

[25] Note that this is consistent with the model, where the labor allocated to a certain region and sector at time t is the deterministic consequence of the distribution of labor and TFP realizations at time $t-1$. Through the lens of the model, the labor allocation L_t^{rs} is exogenous to the actual realization of technological shocks at time t.



(a) Measured TFP in 2012 (b) Stochastic component of TFP in 2012

Figure 6.2: TFP across French regions

isolate the stochastic part and the deterministic part of TFP, we purge TFP of any common time-varying variation in TFP across France (captured by time-fixed effects) as well as the component pertaining to the pre-determined agglomeration of labor:

$$\ln A_t^{rs} = \kappa \ln L_t^{rs} + \delta_r + \delta_s + \delta_t + \eta_t^{rs}, \tag{26}$$

where κ is the elasticity of measured TFP to labor allocation, δ_r, δ_s, and δ_t are, respectively, region-, sector-, and year-fixed effects. The variation in measured TFP explained by these components together is 78.6%. The parameter estimate $\hat{\kappa} = 0.028$ is small but statistically significant. Hence, an agglomeration of labor boosts TFP as in Desmet et al. (2018). Log labor explains somewhat less than 1% of the total variation in log TFP. We define the stochastic component of TFP as the region and sector-specific components of measured TFP plus the residual in (26):

$$\ln a_t^{rs} = \delta_r + \delta_s + \eta_t^{rs}. \tag{27}$$

Panel (b) in Figure 6.2 reports on the distribution of the stochastic component of TFP, a_t^{rs}, across French regions after aggregating using sector-level employment weights. Differences with respect to the overall measured TFP in Panel (a) of the figure are modest in magnitude.[26]

Using the panel data on $\ln a_t^{rs}$ for the years 2003-2014, we postulate and estimate the

[26] Clearly, the variance of the stochastic component a_t^{rs} is smaller than that of overall TFP A_t^{rs}. Hence, by considering only a_t^{rs} to inform the degree of uncertainty, we reduce the scope for uncertainty to play a role relative to a model which would rely on A_t^{rs} altogether.

following autoregressive process that informs the uncertainty process about realized TFP levels:

$$\ln a_t^{rs} = \mu^{rs} + \rho \ln a_{t-1}^{rs} + \iota^{rs} \varepsilon_t^{rs}, \tag{28}$$

where $\varepsilon_t^{rs} \sim i.i.d.(0,1)$, ρ is a common autoregressive coefficient, the region-sector-fixed effect μ^{rs} scales the mean of the process, and ι^{rs} parameterizes the region-sector specific volatility that tunes the degree of uncertainty about region-sector TFP.[27] As long as $|\hat{\rho}| < 1$, the process in (28) is stationary. Then, $\ln a^{rs} \equiv \hat{\mu}^{rs}/(1-\hat{\rho})$ and $\sigma^{rs} \equiv \hat{\iota}^{rs}/\sqrt{1-\hat{\rho}^2}$ are, respectively, the asymptotic long-run mean and standard deviation of the stochastic component of TFP in region-sector rs.

Table 5: Summary statistics for the dynamic process of $\ln(a_t^{rs})$

	mean	std.dev.	p10	p50	p90
Long-run mean $\ln a^{rs}$	2.436	0.108	2.318	2.438	2.543
Long-run std.dev. σ^{rs}	0.056	0.058	0.015	0.037	0.117

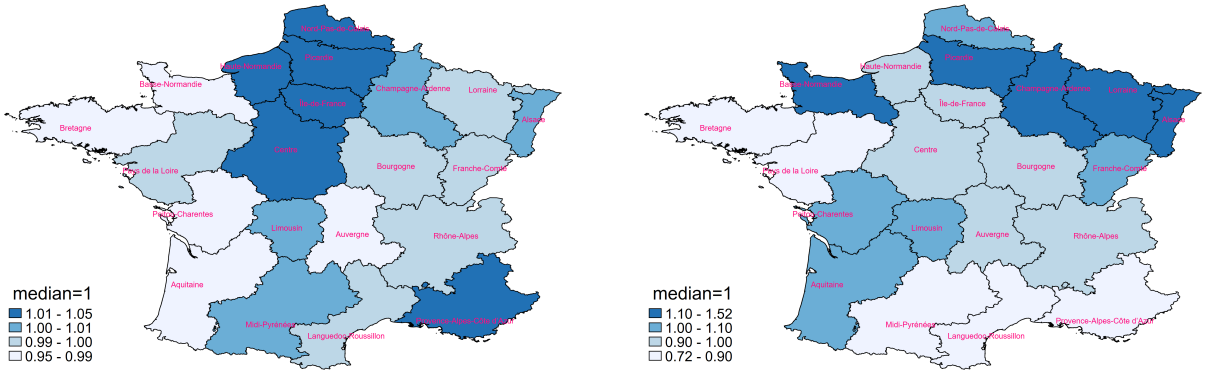
Note: The statistics "mean", "std.dev.", "p10", "p50", and "p90" refer to the average, the standard deviation, and the 10th, 50th, and 90th percentiles of the data across region-sector units. The long-run means and standard deviations in the rows are computed per region-sector unit rs over time.

Table 12 reports on the identified moments of $\ln(a_t^{rs})$, where the estimate of the autoregressive coefficient is $\hat{\rho} = 0.503$, and the volatility $\hat{\iota}^{rs}$ is estimated as the root mean-squared error of the residual based on (28). Panel (a) of Figure 6.3 reports on the distribution of the weighted long-run mean, $\ln a^{rs}$, across French regions, and Panel (b) reports on the weighted long-run standard deviation, σ^{rs}, where weights reflect sector-level employment in each region.

6.3 Calibration and assessment of parameters used in the quantification

Note that the cost-share parameters, trade costs, migration costs, endowments, and TFP levels are derived from real-world data based on equilibrium relationships implied by the model. Following the approach of Dekle et al. (2008), we invert the within-period equilibrium given the observed vector of region-sector-specific wages $\{\bar{w}_{t_0}^{rs}\}_{r=1,s=1}^{R,S}$ and labor allocations

[27] If the panel were very short, we would inherit the well-known Nickell (1981) bias in the estimate of ρ. The latter fades as the number of years grows, and we ignore it here. One could alternatively use instrumental-variation, but the latter also exhibits a small-sample bias.



(a) Long-run mean, $\ln a^{rs}$ (b) Long-run standard deviation, σ^{rs}

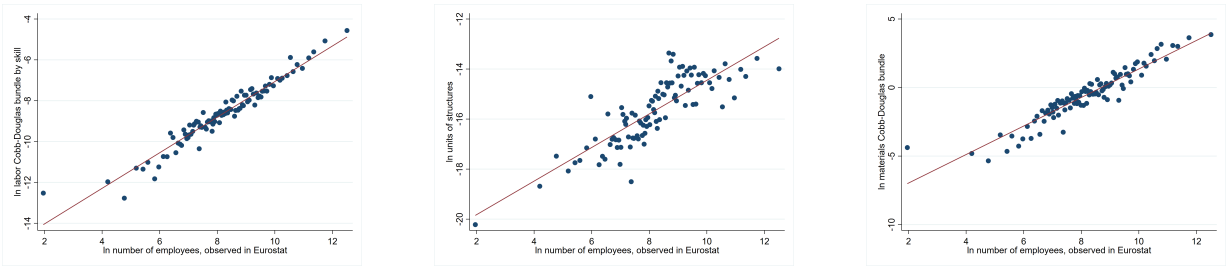
Figure 6.3: Spatial variation of moments of the process for the stochastic component of TFP

$\{L_{t_0}^{rs}\}_{r=1,s=1}^{R,S}$ across jobs in the baseline year $t_0 = 2012$, and given the estimated long-run mean of the stochastic component of TFP levels $\{a^{rs}\}_{r=1,s=1}^{R,S}$, to obtain model-consistent regional trade deficits $\{D_{t_0}^{r}\}_{r=1}^{R}$. This provides an initial equilibrium allocation calibrated to the observed economy in the baseline year 2012.

In order to assess the parameters used to quantify the model, we perform a number of validation exercises. Since the model takes the initial allocation of the economy as given, we want to be sure that units of labor, structures and materials that the calibration implies match the salient features of the French economy in the base year 2012. To this purpose we compare the endowment variables we feed into the model with evidence that we did not use: the number of employees (rather than hours of employment) and the number of firms by sector and region in France, reported by Eurostat in *Regional structural business statistics. Industry, trade and services*.

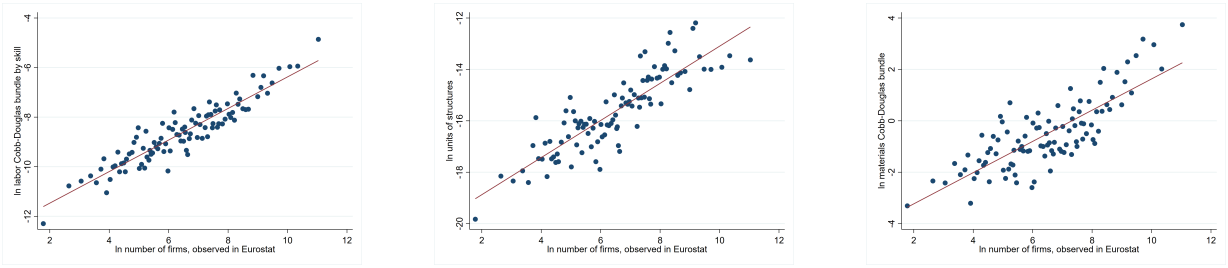
Of course, we expect a high, positive correlation for worker and employee numbers with the hours worked, since we are using hours of labor from matched employer-employee administrative data. But note that what informs the model are units of a Cobb-Douglas bundle of hours of employment of skilled and unskilled workers. Thus, the functional form of production and the skill composition are not necessarily predicting more units of the labor bundle, where we also observe more people employed. A similar argument can be made for materials: the model counts a unit of intermediate inputs as a Cobb-Douglas bundle of products sourced within a region, between regions and internationally, but the spatial correlation with observed economic activity can be checked. For structures, where we rely on Eurostat data on land use and data on fixed assets by industry in the quantification exercise, we can check, whether this metric correlates well with employment and firms present in those regions and

sectors that use this and other factors intensively.



(a) Labor bundle (b) Structures (c) Materials

Figure 6.4: Correlation of model-based quantities with observed number of employees

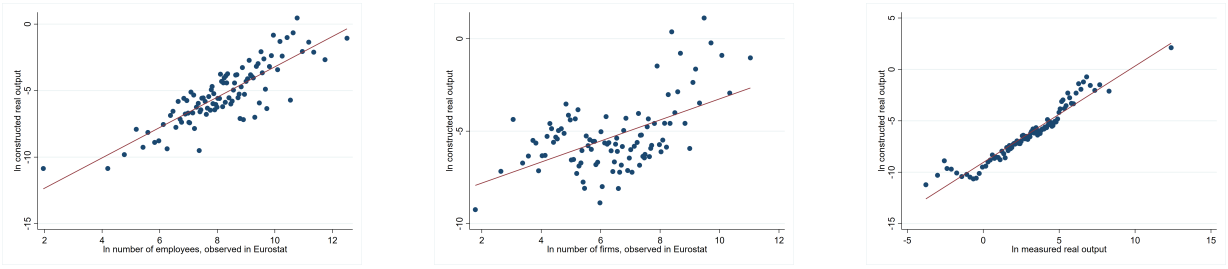


(a) Labor bundle (b) Structures (c) Materials

Figure 6.5: Correlation of model-based quantities with observed number of firms

Figures 6.4 and 6.5 report 100-bin scatter plots of model-based quantities of labor-bundle employment in hours, of structures, and of materials (in logs on the vertical axis, Panels (a), (b) and (c), respectively) versus the number of employees and the number of firms (in logs on the horizontal axis) in 2012. The strong correlations suggest that the calibration of the model is meaningful in matching the observed spatial variation of economic activity in the respective total-employment and firm-numbers domains.

Along the same lines, one can compare how real output by region and sector that is constructed based on Equation (25) using the data for the quantification exercise correlates with observed economic activity obtained from Eurostat data. Figure 6.6 reports 100-bin scatter plots of model-based quantities of output (in log on the vertical axis) constructed using the stochastic component of TFP, a_t^{rs}, as in the quantification instead of measured TFP, A_t^{rs}, versus the observed number of employees and firm numbers (in logs, panel (a) and panel (b), respectively) in 2012. Also in this case, correlations suggest a good match between the spatial distribution of real output used in the quantification of the model and



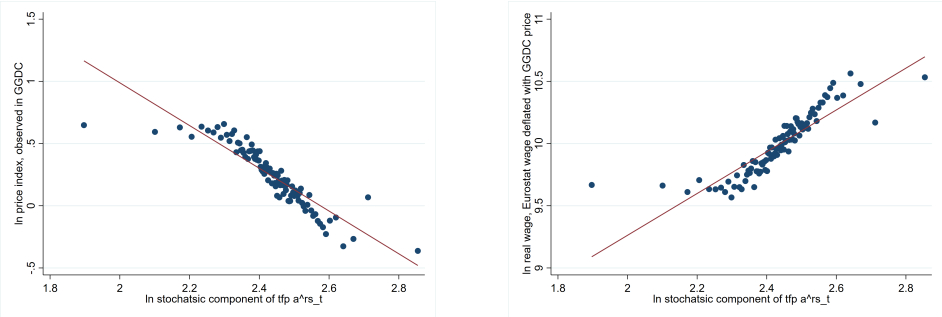
(a) Employees (b) Firms (c) Measured real output

Figure 6.6: Correlation of model-based real output with observed economic activity

the observed spatial variation of economic activity. In Panel (c) we show the correlation between the constructed real output (in logs, vertical axis) and the real output that are consistent with measured TFP (in logs, horizontal axis), in 2012. For the same year, the correlation across region-sector pairs between the two respective measures $\ln a_t^{rs}$ and $\ln A_t^{rs}$ is 49.2%. While the correlation is – not surprisingly – positive, it indicates that the conditional mean of $\ln A_t^{rs}$ accounts for a large share of the variance in $\ln A_t^{rs}$ that is not used in $\ln a_t^{rs}$.

The last evidence is about the relationship between observed real wages and price indices with the estimated stochastic component of TFP, $\ln a_t^{rs}$. The goal is now to gain insights on the implications of the stochastic component of TFP only, rather than of overall measured TFP, for prices and real wages. In the same Eurostat data used for the previous analysis, we observe labor costs at the region-sector level, apart from the number of employees. The ratio of these statistics provides a proxy for the nominal gross annual wage earnings per person employed in a given region and sector. Let us refer to the latter as W_t^{rs}. Region-sector-specific price indices as used in quantifying the model are combined according to the Cobb-Douglas consumption-based price index assumed by the model to obtain a region specific price index; let us refer to this as P_t^r. The ratio W_t^{rs}/P_t^r is a model-consistent measure of observed average real wage in region-sector rs.

The 100-bin scatter plots in Figure 6.7 show the year-2012 correlations of the estimated stochastic component of TFP, $\ln a_t^{rs}$ on the vertical axis, with price indices $\ln P_t^r$, and real wages $\ln W_t^{rs}/P_t^r$ across regions and sectors in Panels (a) and (b), respectively. Region and sector pairs characterized by higher TFP exhibit a lower price index. Real wages are strongly positively correlated with the stochastic component of TFP.



(a) Price index (b) Real wage

Figure 6.7: Correlation of stochastic component of TFP with price index and real wage

7 Quantifying the role of aggregate uncertainty

The focus of the quantification is to illustrate the evolution of the dynamic spatial equilibrium of the model under uncertainty versus perfect foresight for a given initial condition. We will investigate the evolution of welfare and labor allocation over time for an uncertainty scenario and a perfect-foresight scenario. The ideal exercise to quantify a departure of the equilibrium allocation from the perfect foresight due to uncertainty alone is to evaluate the model under both regimes given the same initial allocation at time t_0 and with the same future realizations of economic fundamentals for all periods $t \geq t_0$. By keeping fundamentals constant throughout, dynamics in terms of labor reallocation arise in both scenarios as the initial distribution of labor is not in a steady state, and differences in the evolution of the labor force allocation across the two scenarios emerge solely due to the uncertainty of agents about the realization of TFP in the future. Hence, the only mechanism at work across scenarios is that agents make different optimal forward-looking decisions under uncertainty than under perfect foresight, and this determines a different distribution of labor across jobs.

In what follows, we will use t_0 to indicate the initial period from which we will let the model run forward in both the perfect-foresight and the uncertainty cases. Note that this initial year corresponds almost fully to the estimated model as of 2012. However, there are three exceptions: we assume that each region-sector pair rs realizes from t_0 onwards ad infinitum the long-run TFP level which is consistent with the estimates of the TFP process for each rs; we assume that stochastic shocks around this mean are assumed by the agents to be drawn from the long-run standard deviation with the nodes of a gridded distribution following the Tauchen method under uncertainty as described below;[28] and we calibrate

[28] More precisely, the stochastic component can be drawn arbitrarily from the chosen support. However, we

deficits across regions such that the intratemporal real-wage equilibrium corresponds to the data of 2012 at t_0. None of the fundamental model parameters will change from t_0 onwards, but outcomes may and do adjust during $t > t_0$.

7.1 Counterfactual

In the present context, "*aggregate uncertainty*" materializes through the stochastic component of TFP, i.e., the unexpected part of it. The latter is indexed at the region-sector level. Accordingly, scaling the stochastic component of TFP entails a common shock to all individuals in the same sector and region, irrespective of their age or job. The natural alternative or benchmark to a regime with aggregate uncertainty is one that assumes "*perfect foresight*". Under both regimes, agents have complete knowledge of the model and of the data until t_0. However, under perfect foresight agents have complete and accurate information not only about the past and present but also about future model outcomes and data in $t > t_0$. Instead, under uncertainty agents form expectations based on their knowledge of the stochastic process (28).

7.2 A simple design for modeling aggregate uncertainty

By design, the economy starts at given long-run means for the stochastic component of TFP $a_t^{rs} = a^{rs}$ at $t = t_0$, and it will evolve with zero random innovations, $\varepsilon_t^{rs} = 0$ for every period t and every region-sector pair. This provides a controlled environment in which there is no change in the fundamentals of the economy. We wish to compare predicted outcomes of the model when agents take volatility of the TFP process into account in the *uncertainty regime*, with the *perfect-foresight regime* where agents do not take random shocks to TFP into account, when planning, as the probability of random shocks is zero. Risk-averse agents are rational and "know" the random dynamic process of the stochastic component of TFP in (28). Therefore, the effect of uncertainty is measured by comparing outcomes predicted by the model informed with the estimated long-run mean and volatility parameters with outcomes predicted by the model given zero volatility.

The economy is characterized by a vector of $(R + 1) \cdot S$ stationary random processes, thus, the parameters σ^{rs} are the (squared) diagonal elements of the covariance matrix of a vector-autoregressive process (VAR) with region-sector-specific innovations. By modeling

pick ex post exactly that draw, which corresponds to the long-run mean. Hence, the realization of TFP is fixed at the long-run mean in every period $t \geq t_0$ also under uncertainty. But the agents do not foresee this.

and estimating the covariance matrix, one could disentangle the role of *local* uncertainty with $\sigma^{rs} > 0$ for a given region-sector pair (rs) but $\sigma^{r's'} = 0$ for all $(r's') \neq (rs)$, or *global* uncertainty with $\sigma^{rs} > 0$ in all region-sector pairs jointly. One could also design experiments that combine these extremes. We focus on the case of global uncertainty with the world being in one of three states of nature.[29] Note that this stochastic process describes the stochastic component of TFP, i.e., a_t^{rs}, whereas measured productivity, i.e., A_t^{rs}, will in general exhibit spatial and time correlation patterns as an endogenous outcome of the model, a consequence of the deterministic labor reallocation across jobs over time.

Table 6: Nodes of the Markov chain for a_t^{rs}

	node 1	node 2	node 3
relative to median	0.866	1	1.144
median	9.917	11.453	13.097
std.dev.	2.098	1.219	2.140

Table 7: Transition probabilities

from \to	node 1	node 2	node 3
node 1	0.5003	0.4994	0.0003
node 2	0.0416	0.9168	0.0416
node 3	0.0003	0.4994	0.5003

We proceed with a parsimonious configuration that approximates the $(R+1)\cdot S$ processes in (28) with a Markov chain for the standardized moments using the Tauchen (1986) method with $Q = 3$ nodes for the aggregate stochastic state of TFP. Table 6 reports the median and standard deviation across region-sector pairs within each of the three nodes of the Markov chain. The first row of the table reports on the median within each node relative to the median of the central node. Table 7 summarizes the transition probabilities between the nodes of the chain. Nodes for the stochastic component of TFP are within a range of 15% around the long-run mean, and agents attach a probability of 91.68% to a stochastic component of TFP that remains at its long-run mean, i.e., $a_{t+1}^{rs} = a^{rs}$ conditional on $a_t^{rs} = a^{rs}$. The associated transition probabilities of stochastic TFP between the considered nodes generate a stochastic TFP support that is centered without large deviations around expected

[29] More sophisticated modeling choices remain tractable. The "vector case" in Tauchen (1986) accommodates vector autoregressions in which realizations of the stochastic component in a region and sector, $\ln a_t^{rs}$, are a random function of previous own levels $\ln a_{t-1}^{rs}$ but also previous levels in other region-sector pairs, e.g., $\ln a_{t-1}^{r's}$ or $\ln a_{t-1}^{rs'}$ or $\ln a_{t-1}^{r's'}$. The extension of these methods developed by Terry and Knotek (2011) accommodates the case of arbitrary positive-semidefinite covariance structures of the innovations. These methodologies can be computationally intense for a large number of region-sector pairs, but recent developments are useful to conduct efficient VAR discretizations, that also accommodate granularity typical of joint probabilities on large dimensional stochastic spaces; for instance, see Gordon (2021). An assessment of those techniques is out of the scope of this paper, but all of them provide a quite accurate approximation of first-order VARs with transition probabilities on a finite (reasonably small) number of realizations on the stochastic support; respectively, $\pi_{q,q'}$ and z^q for $q, q' = 1, ..., Q$ in the notation of Section 5.

TFP, that coincides with the perfect-foresight scalar-valued counterpart for each region and sector.

Overall, this level of uncertainty can be considered relatively conservative compared to perfect foresight. Agents not only correctly predict the expected value of TFP but also anticipate its reasonably small variance in the exercise.

7.3 Distribution of labor across jobs

The distribution of labor across jobs is the aggregate deterministic state of the MFG, and in the initial allocation in period t_0 it is given by the observed distribution. Additionally, we have to select alternative distributions of labor across jobs for interpolating the payoffs associated with forward-looking decisions under uncertainty as described in Section 5.3.

A choice for the basis of distributions should be such that the solution of the model (for the initial period t_0 and for all periods thereafter) likely belongs to its convex hull. Given this guideline, we consider $P = 5$ distributions: l^1 is the observed distribution of labor across jobs in 2012, and the remaining ones are implied by the solution of the myopic DOCP in which agents make their decisions, assuming that the next-period distribution is equal to the current one.[30] Call $\Delta L = \tilde{L}_1 - L_0$ the difference between the one-period-ahead solution of the myopic problem and the initial allocation. Then, $l^2 = L_0 + \Delta L$ is the one-period-ahead myopic solution and $l^3 = L_0 - \Delta L$ is the distribution obtained by computing changes from the myopic solution in the opposite direction. Finally, distributions $l^4 = \tilde{M}^{25} L_0$ and $l^5 = \tilde{M}^{50} L_0$ are, respectively, the 25th- and 50th-period-ahead solutions of the myopic problem with $\tilde{M}$ being the one-period policy of the myopic DOCP.

This setup is sufficiently rich to compute the numerical solution of the MFG with accuracy, and it is sufficiently parsimonious to illustrate the solution algorithm. As discussed in Section 5, the numerical solution is obtained by iteration on the system of $P \cdot Q = 15$ MFGs, each with an initial allocation (l^p, z^q) for $p = 1, ..., 5$ and $q = 1, ..., 3$. We start by guessing that the next-period distribution of labor across jobs is equal to the current distribution. Then, the algorithm proceeds as follows. At each iteration, the next-period distribution implied by the Kolmogorov operator given the current guess of value functions is used in the next iteration to interpolate the updated guess of value functions. Table 8 reports the

[30] The myopic problem is not an MFG, since it consists only of the Bellman equation, and it is not an equilibrium, since the next-period distribution implied by the aggregation of individual decisions does not correspond to the current distribution, as assumed by the agents. However, the myopic problem still contains information about all the economic fundamentals of the MFG (e.g., preferences, technologies, frictions), and it can be solved directly by value-function iteration on the Bellman equation.

average barycentric weights of each distribution (listed on the columns) over the iterations. This is done for each initial labor allocation in Panels (a)-(e) and for each realization of the stochastic state in the rows. Barycentric weights for a next-period distribution L implied by the Kolmogorov equation are constructed by first selecting the two distributions among $(l^1, ..., l^P)$, whose convex hull contains the largest fraction of L, and then computing the weights using uniform distance.

The structure that emerges from this procedure is one of a network. For the solution of the MFGs with given initial labor allocation $L_0 = l^1$ (that is the distribution of labor across jobs in France in 2012, here), the algorithm interpolates on the basis l^1 and l^4 predominantly (i.e., across the considered stochastic states of TFP), while l^2 contributes only when the initial aggregate stochastic state is $Z_0 = z^2$. The basis distributions l^3 and l^5 are never selected as to be the closer ones to the solution of the Kolmogorov equation. However, to compute the solution for the value of MFGs with given initial labor allocation $L_0 = l^4$, the algorithm selects the basis l^5 (together with l^1, l^2 and l^4). And to compute the solution of MFGs with given initial labor allocation $L_0 = l^5$, the algorithm selects the basis l^3 as well. On a customary laptop, the algorithm computing the solution of the system of 15 MFGs converges after 133 iterations in less than 500 seconds to a solution which exhibits a tolerance of less than 10^{-5} for the sup distance of the vector value functions (whose average is 88.19) and less than 10^{-9} for the sup distance on the labor distribution (whose average is 4×10^{-5}).

This illustrates how to obtain the solution to the numerical MFG (14)-(15), given L_0. With the distribution L_1 predicted by the model, the piece-wise linear interpolation (18)-(20) can be computed to evaluate the evolution of the MFG, hence, L_2, and so forth. This allows us to evaluate the transitional dynamics of the model and deal with patterns that are due to the initial allocation not being a steady state. Specifically, to let the memory of the initial allocation vanish, we simulate 100 years after 2012. Despite the parsimonious setup, the approximation errors on both the Bellman operator (23) and the Kolmogorov operator (24) remain below 1% over the entire simulated period. With this level of accuracy, we proceed and examine the results.

7.4 Results

In presenting the results of the quantification exercise, we start from the main message and then explain the mechanisms at work. We initialize the model with the allocation of labor across regions, sectors, and occupations in France as of 2012 and with the stochastic component of TFP pertaining to the region-sector-specific long-run mean. We then solve

Table 8: Average barycentric weights across iterations

	l^1	l^2	l^3	l^4	l^5
initial aggregate stochastic state is $Z_0 = z^1$	0.9028	0	0	0.0972	0
initial aggregate stochastic state is $Z_0 = z^2$	0.8962	0.0075	0	0.0963	0
initial aggregate stochastic state is $Z_0 = z^3$	0.9028	0	0	0.0972	0

a. initial labor allocation is $L_0 = l^1$

	l^1	l^2	l^3	l^4	l^5
initial aggregate stochastic state is $Z_0 = z^1$	0.7943	0.0066	0	0.1991	0
initial aggregate stochastic state is $Z_0 = z^2$	0.7878	0.0141	0	0.1982	0
initial aggregate stochastic state is $Z_0 = z^3$	0.7943	0.0066	0	0.1991	0

b. initial labor allocation is $L_0 = l^2$

	l^1	l^2	l^3	l^4	l^5
initial aggregate stochastic state is $Z_0 = z^1$	0.0066	0.5609	0.4317	0.0009	0
initial aggregate stochastic state is $Z_0 = z^2$	0	0.5683	0.4317	0	0
initial aggregate stochastic state is $Z_0 = z^3$	0.0066	0.5609	0.4317	0.0009	0

c. initial labor allocation is $L_0 = l^3$

	l^1	l^2	l^3	l^4	l^5
initial aggregate stochastic state is $Z_0 = z^1$	0.0066	0.3218	0	0.0009	0.6707
initial aggregate stochastic state is $Z_0 = z^2$	0	0.3293	0	0	0.6707
initial aggregate stochastic state is $Z_0 = z^3$	0.0066	0.3218	0	0.0009	0.6707

d. initial labor allocation is $L_0 = l^4$

	l^1	l^2	l^3	l^4	l^5
initial aggregate stochastic state is $Z_0 = z^1$	0.0066	0.5106	0.4819	0.0009	0
initial aggregate stochastic state is $Z_0 = z^2$	0	0.5181	0.4819	0	0
initial aggregate stochastic state is $Z_0 = z^3$	0.0066	0.5106	0.4819	0.0009	0

e. initial labor allocation is $L_0 = l^5$

Note: This table reports average barycentric weights for the $P = 5$ distributions of jobs and $Q = 3$ realizations of the aggregate stochastic state that serve as the basis for discretization.

the model under two alternative regimes: under *aggregate uncertainty*, with the matrix of transition probabilities between aggregate TFP states as given in Table 7, and under *perfect*

foresight, where agents (correctly) attach a probability of one to the event of a stochastic component of TFP equal to its long-run mean everywhere. For all scenarios, we assume agents' relative-risk-aversion parameter to be $\chi = 1$ which corresponds to log utility.

Tables 9 and 10 report moments of the distribution of percentage changes in lifetime values of all jobs between the two regimes for agents in the old age spell and the young age spell, respectively. A negative sign indicates a lower value associated with the uncertainty regime. The numbers in the table suggest that, on average, uncertainty entails a welfare loss: evaluated in t_0, the average present-discounted lifetime value of a job falls by -0.20% for an old agent and by -0.44% for a young agent due to uncertainty alone. The aggregate welfare loss is due to two important forces. First, for risk-averse agents the chance of negative productivity shocks entails welfare losses. Second, akin to the effect uncertainty has in other macroeconomic contexts such as investment decisions, arbitrage mechanisms across jobs are less effective because of a slower reallocation which quasi imprisons agents in bad locations. However, a rich set of findings emerges regarding the heterogeneity of the impact of uncertainty across jobs and over time.

Table 9: Change in lifetime value of a job, uncertainty minus perfect foresight in percent for the old age spell

	mean	std.dev.	p1	p10	p50	p90	p99
t_o	-0.20	3.77	-5.456	-0.563	-0.029	0.859	3.829
10 years after	-0.19	3.72	-5.67	-0.762	-0.022	0.93	4.062
50 years after	-0.22	3.78	-6.80	-1.316	-0.013	1.288	5.119
100 years after	-0.28	3.93	-7.94	-2.017	-0.010	1.543	6.463

Table 10: Change in lifetime value of a job, uncertainty minus perfect foresight in percent for the young age spell

	mean	std.dev.	p1	p10	p50	p90	p99
t_o	-0.44	5.31	-17.034	-0.964	-0.031	1.198	4.368
10 years after	-0.44	5.30	-17.019	-1.126	-0.025	1.231	4.560
50 years after	-0.46	5.30	-17.024	-1.455	-0.019	1.498	5.190
100 years after	-0.50	5.32	-17.086	-1.967	-0.018	1.782	5.821

Importantly, not all jobs yield lower present-discounted lifetime values, but some become more attractive compared to the perfect-foresight scenario and are then comparatively better off. Hence, there are comparative winners from uncertainty. Evaluated at t_0, the old-age-spell uncertainty-to-perfect-foresight welfare differences range from -5.46% to $+3.83\%$. The

young-age-spell support of differences is even bigger and ranges from -17.03% to $+4.37\%$. In a spatial setting as considered here, this heterogeneity in the welfare effects leads to a different allocation of labor across space even with an identical realization of fundamentals, hence, purely because of uncertainty. To be more precise: (i) differences in fundamentals (structure endowments, mobility and trade costs, etc.) translate an identical degree of uncertainty across jobs into heterogeneous welfare effects and associated labor-distribution responses with winners and losers (due to response heterogeneity); (ii) differences in uncertainty lead to heterogeneous welfare effects and associated labor-distribution responses with winners and losers (due to treatment heterogeneity); (iii) and differences in uncertainty as well as fundamentals lead to heterogeneous welfare effects and associated labor-distribution responses with winners and losers (due to an interaction of treatment and response heterogeneity).

In an aggregate macroeconomic environment without multiple spatial units a state is either preferable or not. In such a setting, *relatively low uncertainty* does not emerge as a source of *comparative advantage* which can be preferable to a situation with perfect foresight for some jobs. With agents disliking uncertainty, a scenario with perfect foresight would be preferable. But with spatial units that are connected via mobility (as well as trade), the notion of an absolute welfare loss from uncertainty is accompanied by one of *relative uncertainty* as a source of *comparative advantage* (in its interaction with deterministic fundamentals).

Here, each spatial unit is affected differently by uncertainty (treatment heterogeneity), and the equilibrium that links all spatial units induces complex interactions of this uncertainty in terms of its implication for labor reallocation (response heterogeneity). The two types of heterogeneity lead to reallocations that can ceteris paribus make agents in some jobs even better off on average under uncertainty than under perfect foresight. The quantitative difference in the attractiveness of jobs between uncertainty and perfect foresight is huge here, as can be seen from the standard deviation in welfare differences across jobs being one order of magnitude larger than the mean. This implies that the ranking of jobs in terms of their attractiveness and their relative differences changes in important ways. Given the same level of moving costs and idiosyncratic preferences, this will lead to different optimal location decisions under the two regimes even at the same level of fundamentals in all places. In this sense, uncertainty shapes the spatial economy.

One way to see how uncertainty shapes the spatial economy is by looking at Figure 7.1 where we present the distribution of each job's size under uncertainty relative to perfect foresight after 100 years. A substantial share of jobs is either substantially smaller or larger compared to a perfect-foresight scenario in terms of employment. These differences emerge

due to the different attractiveness of jobs across scenarios as illustrated in Figure 7.2 which presents a scatter plot of the rank of jobs in terms of their lifetime value across scenarios for old (left panel) and young (right panel), respectively.

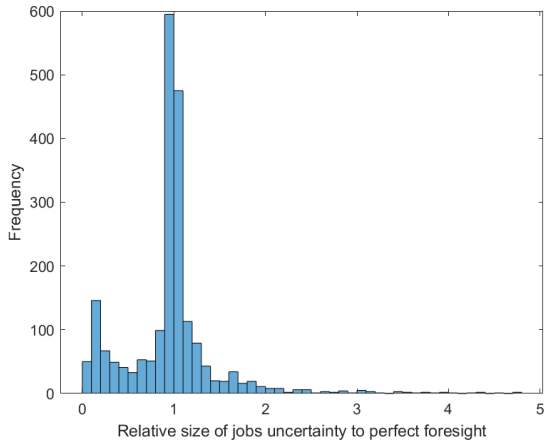
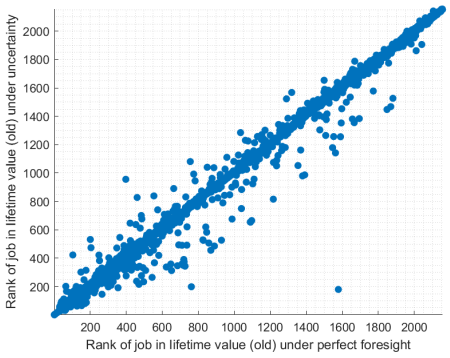
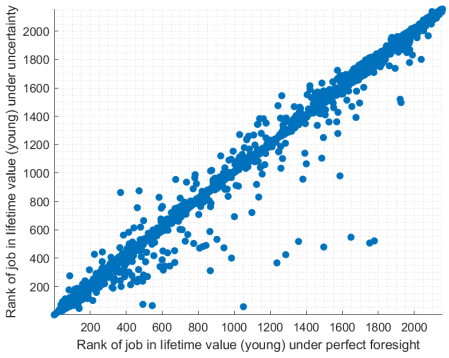


Figure 7.1: Size of jobs under uncertainty and perfect foresight.



Scatter plot of job rank
in terms of lifetime value for old



Scatter plot of job rank
in terms of lifetime value for young

Figure 7.2: Ranking of jobs under uncertainty and perfect foresight.

Tables 9 and 10 show that the patterns of uncertainty-to-certainty differences are persistent and do not vanish with labor reallocation. Even 100 years after t_0, the distribution of attractiveness of jobs, measured by the average discounted life-time value of being there, is substantially different between perfect foresight and uncertainty.

Moreover, a comparison of Tables 9 and 10 as well as a comparison of the two panels in Figure 7.2 shows that differences in the lifetime horizon matter. The outcomes of choices made under uncertainty deviate from their perfect-foresight analogue substantially more for

young than for old agents. Given that utility and moving costs are the same in the two age spells, the source of this differential impact is entirely due to the difference in continuation values. The latter differ by age, because young agents have a longer lifetime horizon than old ones, whereby the discounted continuation value is larger for the young than for the old. Therefore, old agents have a ceteris paribus lower incentive to reallocate because of a shorter lifetime horizon (in which moving costs are less likely balanced by the benefit of escaping uncertainty). Conversely, for young agents reallocation pays off more easily than for old ones.

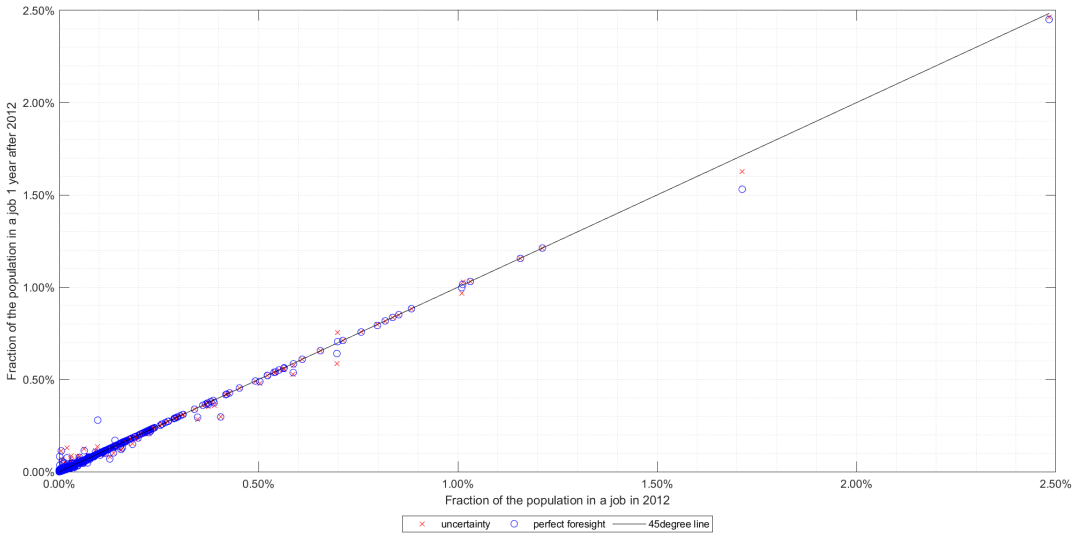


Figure 7.3: Labor reallocation in t_o, uncertainty versus perfect foresight

Given the initial allocation of labor, the model predicts a fraction of old agents that within a year do not change their job of 99.31% and among young agents the same statistics fall to 97.40%. This matches the evidence: in the data, the average fraction of agents that within a year do not change job is 98.20%, and in the model the weighted average across age spells is 98.54%. The difference between old and young cohorts in terms of moving despite identical moving costs is explained by the differential value that agents attach to reallocation depending on their lifetime horizon.

These insights show that uncertainty alone changes the value of jobs in heterogeneous and complex ways that ultimately determine a different allocation of *who* produces *what* and *where*. The long-run allocation of production and consumption emerges through reallocation of labor starting from the same initial allocation. But reallocation under uncertainty and

perfect foresight differ due to the aforementioned mechanisms together at work. Importantly, as the initial state of the economy is not necessarily a steady state for any of the scenarios, we can compare the differential evolution of the distribution of labor across jobs.

In Figure 7.3, we plot the observed allocation of the labor force across jobs in 2012 (horizontal axis) against the predicted allocation one year after (vertical axis), under uncertainty (jobs are marked with a cross) and under perfect foresight (jobs are marked with a circle). If the initial allocation had been a steady state and in the absence of any shocks, jobs would line up along the 45-degree line. In the chosen setup, we observe transitional dynamics under the uncertainty as well as the perfect-foresight regime. Furthermore, even after one period, differences in the optimal labor allocation are visible. Workers move out of jobs with the higher initial shares under both regimes, but there is more inertia under uncertainty. At the same time, bigger changes are observed for jobs that are less populated, and these changes tend to be rather different across regimes.

Figure 7.4 produces the same information but for the longer run, 100 years after t_0. At this point, the equilibrium has lost memory of the initial allocation, and the economy is characterized by a substantially different allocation of labor across jobs in comparison to t_0. Moreover, the predictions are systematically different between the uncertainty and the perfect-foresight regimes. Hence, the long-run spatial economy is clearly very different under a regime of uncertainty from the one under perfect foresight.

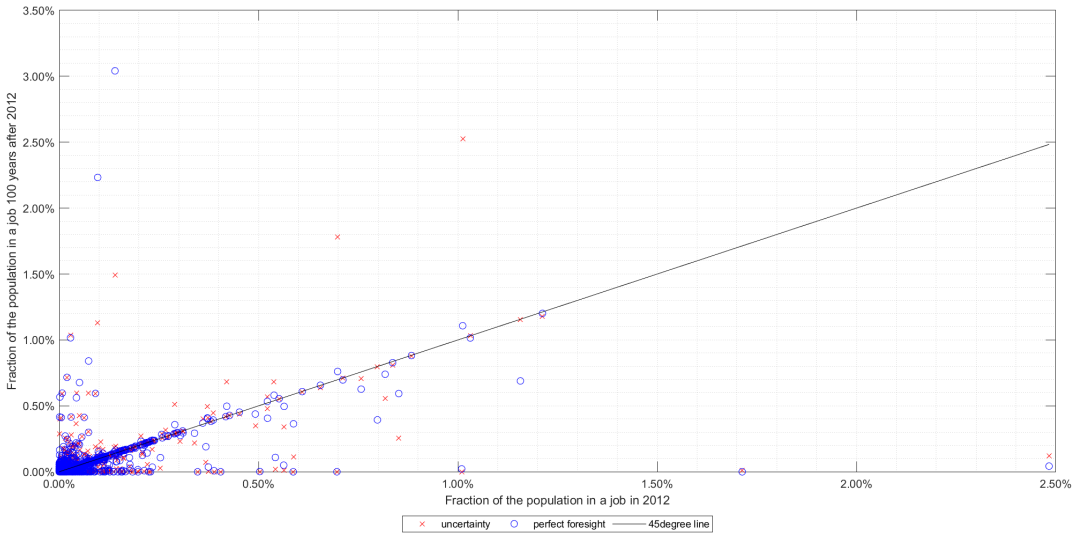


Figure 7.4: Labor reallocation in the long run ($t_o + 100$), uncertainty versus perfect foresight

8 Conclusions

This paper illustrates a novel methodology that allows us to consider uncertainty about future aggregate outcomes, designed in discrete-time and discrete state space to be specifically suited for quantitative dynamic spatial general-equilibrium models. Thanks to this, we document how the distribution of labor across jobs, hence of economic activity across locations and sectors, differs systematically from what is predicted under a perfect-foresight solution.

We directly speak to the concern that people – both as workers and consumers – tend to be risk-averse. When prompted to change jobs and locations, they compare future uncertain returns to their current residence and occupational choices. Thus, in addition to the "freezing role" documented in macro literature regarding dynamic choices, uncertainty plays a broader role in a spatial context. By distorting spatial reallocation compared to perfect foresight, uncertainty creates comparative advantages or disadvantages across jobs by making the opportunity value of reallocating to some jobs greater or lower than under perfect foresight.

To understand and quantify how aggregate uncertainty affects the evolution of the spatial distribution of people and economic activity, we quantify a dynamic stochastic general-equilibrium open-economy model with interregional and intersectoral movements of goods and people, who also might choose at any period to change occupation, sector or location. Within each period, the economy works as a state-of-the-art Eaton-Kortum-type production equilibrium with input-output linkages as developed by Caliendo et al. (2019). The specific feature of our framework relies on the intertemporal allocation of labor: Worker-consumer agents make decisions on jobs and locations under uncertainty about productivity changes happening worldwide, being aware of their finite lifetime. By considering rational, aging, forward-looking, and risk-averse households in a dynamic stochastic general equilibrium model with realistic spatial linkages due to the moving costs of both goods and people, our theory provides a rich understanding of the way in which uncertainty impacts reallocation in the spatial economy.

A large-scale quantification conducted on France demonstrates systematic and significant changes in the relative attractiveness of jobs solely due to uncertainty. Given the different allocation of labor across locations and sectors, sizeable welfare losses, but also gains, emerge under uncertainty compared to perfect foresight.

References

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from us labor markets. Journal of Political Economy, 128(6):2188–2244.

Achdou, Y. and Capuzzo-Dolcetta, I. (2010). Mean field games: Numerical methods. SIAM Journal on Numerical Analysis, 48(3):1136–1162.

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2021). Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. The Review of Economic Studies, 89(1):45–86.

Adao, R., Arkolakis, C., and Esposito, F. (2020). General Equilibrium Effects in Space: Theory and Measurement. NBER Working Papers 25544, National Bureau of Economic Research, Inc.

Allen, T. and Donaldson, D. (2020). Persistence and path dependence in the spatial economy. Working Paper 28059, National Bureau of Economic Research.

Alvarez, F. and Lucas, R. E. (2007). General equilibrium analysis of the eatonÃ¢Â€Â“kortum model of international trade. Journal of Monetary Economics, 54(6):1726–1768.

Artuç, E., Chaudhuri, S., and McLaren, J. (2010). Trade shocks and labor adjustment: A structural empirical approach. American Economic Review, 100(3):1008–45.

Artuç, E. and McLaren, J. (2015). Trade policy and wage inequality: A structural analysis with occupational and sectoral mobility. Journal of International Economics, 97(2):278–294.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. Econometrica, 89(5):2375–2408.

Baker, S. R., Bloom, N., and Terry, S. J. (2023). Using Disasters to Estimate the Impact of Uncertainty. The Review of Economic Studies, 91(2):720–747.

Bilal, A. (2023). Solving heterogeneous agent models with the master equation. Working Paper 31103, National Bureau of Economic Research.

Bilal, A. and Rossi-Hansberg, E. (2021). Location as an Asset. Working Paper 24867, National Bureau of Economic Research.

Bilal, A. and Rossi-Hansberg, E. (2023). Anticipating climate change across the united states. Working Paper 31323, National Bureau of Economic Research.

Bloom, N. (2009). The impact of uncertainty shocks. Econometrica, 77(3):623–685.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., and Terry, S. J. (2018). Really uncertain business cycles. Econometrica, 86(3):1031–1065.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. Econometrica, 87(3):741–835.

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of nafta. The Review of Economic Studies, 82(1 (290)):1–44.

Caliendo, L., Parro, F., Rossi-Hansberg, E., and Sarte, P.-D. (2018). The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy. Review of Economic Studies, 85(4):2042–2096.

Debrabant, K. and Jakobsen, E. R. (2013). Semi-lagrangian schemes for linear and fully non-linear diffusion equations. Mathematics of Computation, 82(283):1433–1462.

Dekle, R., Eaton, J., and Kortum, S. (2007). Unbalanced trade. American Economic Review, 97(2):351–355.

Dekle, R., Eaton, J., and Kortum, S. (2008). Global rebalancing with gravity: Measuring the burden of adjustment. Technical Report 55, IMF.

Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2018). The geography of development. Journal of Political Economy, 126(3):903–983.

Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. Econometrica, 82(3):825–885.

Dix-Carneiro, R. and Kovak, B. K. (2017). Trade liberalization and regional dynamics. American Economic Review, 107(10):2908–46.

Dix-Carneiro, R., Pessoa, J. P., Reyes-Heroles, R., and Traiberman, S. (2023). Globalization, Trade Imbalances, and Labor Market Adjustment*. The Quarterly Journal of Economics, 138(2):1109–1171.

Dix-Carneiro, R., Pessoa, J. P., Reyes-Heroles, R. M., and Traiberman, S. (2021). Globalization, trade imbalances and labor market adjustment. Working Paper 28315, National Bureau of Economic Research.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. Econometrica, 70(5):1741–1779.

Fan, J., Hong, S., and Parro, F. (2023). Learning and expectations in dynamic spatial economies. Working Paper 31504, National Bureau of Economic Research.

Fernandez-Villaverde, J., Guerron-Quintana, P., Rubio-Ramirez, J. F., and Uribe, M. (2011). Risk matters: The real effects of volatility shocks. American Economic Review, 101(6):2530–61.

Gomes, D. A., Mohr, J., and Souza, R. R. (2010). Discrete time, finite state space mean field games. Journal de MathÃ©matiques Pures et AppliquÃ©es, 93(3):308–328.

Gordon, G. (2021). Efficient var discretization. Economics Letters, 204:109872.

Han, J., Yang, Y., and E, W. (2022). Deepham: A global solution method for heterogeneous agent models with aggregate shocks. Technical Report 3990409.

Handley, K. and Limao, N. (2017). Policy uncertainty, trade, and welfare: Theory and evidence for china and the united states. American Economic Review, 107(9):2731–83.

Head, K. and Ries, J. (2001). Increasing returns versus national product differentiation as an explanation for the pattern of u.s.-canada trade. American Economic Review, 91(4):858–876.

Heathcote, J., Storesletten, K., and Violante, G. L. (2009). Quantitative macroeconomics with heterogeneous households. Annual Review of Economics, 1(1):319–354.

Kaplan, G. and Violante, G. L. (2018). Microeconomic heterogeneity and macroeconomic shocks. Journal of Economic Perspectives, 32(3):167–94.

Kleinman, B., Liu, E., and Redding, S. J. (2021). Dynamic Spatial General Equilibrium. NBER Working Papers 29101, National Bureau of Economic Research, Inc.

Krishna, P., Poole, J. P., and Senses, M. Z. (2012). Trade, labor market frictions, and residual wage inequality across worker groups. American Economic Review, 102(3):417–23.

Krishna, P. and Senses, M. Z. (2014). International trade and labour income risk in the u.s. The Review of Economic Studies, 81(1 (286)):186–218.

Krusell, P. and Smith, Jr., A. (1998). Income and wealth heterogeneity in the macroeconomy. Journal of Political Economy, 106(5):867–896.

Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. American Economic Review, 108(12):3855–90.

Moretti, E. (2011). Local labor markets. volume 4B, chapter 14, pages 1237–1313. Elsevier, 1 edition.

Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica, 49(6):1417–1426.

Pierce, J. R. and Schott, P. K. (2016). The surprisingly swift decline of us manufacturing employment. American Economic Review, 106(7):1632–62.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. Annual Review of Economics, 9(1):21–58.

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. Journal of Economic Dynamics and Control, 33(3):649–665.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. Economics Letters, 20(2):177–181.

Terry, S. J. and Knotek, E. S. (2011). Markov-chain approximations of vector autoregressions: Application of general multivariate-normal integration techniques. Economics Letters, 110(1):4–6.

Appendices (For Online Publication)

A Dynamic discrete-choice optimal control problem

Rearranging the Bellman equations (1) yields

$$\begin{aligned} W^{bj}\left(L_t, Z_t, \varepsilon_{h,t}\right) &= \beta \mathbb{E}_t\left[V^{yj}\left(L_{t+1}, s_{t+1}\right)\right] + \nu \max_{n \in J}\left\{\varepsilon_{h,t}^{n} + \bar{\varepsilon}_t^{bj,n}\right\}, \\ W^{yj}\left(L_t, Z_t, \varepsilon_{h,t}\right) &= u(c_t^j) + \beta \mathbb{E}_t\left[\bar{V}^{yj}\left(L_{t+1}, s_{t+1}\right)\right] + \nu \max_{n \in J}\left\{\varepsilon_{h,t}^{n} + \bar{\varepsilon}_t^{yj,n}\right\}, \\ W^{oj}\left(L_t, Z_t, \varepsilon_{h,t}\right) &= u(c_t^j) + \beta \mathbb{E}_t\left[\bar{V}^{oj}\left(L_{t+1}, s_{t+1}\right)\right] + \nu \max_{n \in J}\left\{\varepsilon_{h,t}^{n} + \bar{\varepsilon}_t^{oj,n}\right\}, \end{aligned}$$

where

$$\begin{aligned} \bar{\varepsilon}_t^{bj,n} &\equiv \frac{1}{\nu}\left(\beta \mathbb{E}_t\left[V^{yn}\left(L_{t+1}, s_{t+1}\right) - V^{yj}\left(L_{t+1}, s_{t+1}\right)\right] - \zeta^{bj,n}\right), \\ \bar{\varepsilon}_t^{yj,n} &\equiv \frac{1}{\nu}\left(\beta \mathbb{E}_t\left[\bar{V}^{yn}\left(L_{t+1}, s_{t+1}\right) - \bar{V}^{yj}\left(L_{t+1}, s_{t+1}\right)\right] - \zeta^{yj,n}\right), \\ \bar{\varepsilon}_t^{oj,n} &\equiv \frac{1}{\nu}\left(\beta \mathbb{E}_t\left[\bar{V}^{on}\left(L_{t+1}, s_{t+1}\right) - \bar{V}^{oj}\left(L_{t+1}, s_{t+1}\right)\right] - \zeta^{oj,n}\right) . \end{aligned} \tag{29}$$

Suppose that for a given household h the new job which maximizes the expected future value starting from j is i, so that $i = \arg\max_{n \in J}\left\{\varepsilon_{h,t}^{n} + \bar{\varepsilon}_t^{aj,n}\right\}$ and $\max_{n \in J}\left\{\varepsilon_{h,t}^{n} + \bar{\varepsilon}_t^{aj,n}\right\} = \varepsilon_{h,t}^{i} + \bar{\varepsilon}_t^{aj,i}$. This is true if and only if $\varepsilon_{h,t}^{i} + \bar{\varepsilon}_t^{aj,i} \geq \varepsilon_{h,t}^{m} + \bar{\varepsilon}_t^{aj,m}$ for every $m \neq i$.

Let $F(\varepsilon)$ be the c.d.f. of the i.i.d. idiosyncratic shocks $\varepsilon_{h,t}$ and $f(\varepsilon)$ be the corresponding p.d.f., respectively. Realizations of idiosyncratic values are assumed to be independent across jobs, and this implies that the event $\varepsilon_{h,t}^{m} \leq \varepsilon_{h,t}^{i} + \bar{\varepsilon}_t^{aj,i} - \bar{\varepsilon}_t^{aj,m}$ for every $m \neq i$ occurs with the joint probability density $f(\varepsilon_{h,t}^{i}) \prod_{m \neq i} F(\varepsilon_{h,t}^{i} + \bar{\varepsilon}_t^{aj,i} - \bar{\varepsilon}_t^{aj,m})$. Taking the expectation of $\varepsilon_{h,t}^{i}$ over the continuum of households yields $\int_{-\infty}^{\infty}(\varepsilon + \bar{\varepsilon}_t^{aj,i}) f(\varepsilon) \prod_{m \neq i} F(\varepsilon + \bar{\varepsilon}_t^{aj,i} - \bar{\varepsilon}_t^{aj,m}) d\varepsilon$, which is the contribution of a single job to the value of $\mathbb{E}_h\left[\max_{n \in J}\left\{\varepsilon_{h,t}^{n} + \bar{\varepsilon}_t^{aj,n}\right\}\right]$. Summing over all candidate jobs yields the average option value among households of age a for choosing the next-period job starting from being at job j in the current period t

$$\Omega(\bar{\varepsilon}_t^{aj}) \equiv \sum_{i \in J} \int_{-\infty}^{\infty} (\varepsilon + \bar{\varepsilon}_t^{aj,i}) f(\varepsilon) \prod_{m \neq i} F(\varepsilon + \bar{\varepsilon}_t^{aj,i} - \bar{\varepsilon}_t^{aj,m}) d\varepsilon \ , \quad a = \{b, y, o\} \ .$$

Therefore, the average option value of moving at age $a \in \{b, y, o\}$ when starting from a job j is given by

$$\nu\Omega(\bar{\varepsilon}_t^{aj}) \equiv \sum_{i \in J} \int_{-\infty}^{\infty} (\nu\varepsilon + \nu\bar{\varepsilon}_t^{aj,i}) f(\varepsilon) \prod_{m \neq i} F(\varepsilon + \bar{\varepsilon}_t^{aj,i} - \bar{\varepsilon}_t^{aj,m}) d\varepsilon \ , \tag{30}$$

where

$$\nu\bar{\varepsilon}_t^{aj,n} \equiv \beta\mathbb{E}_t\left[\bar{V}^{an}\left(L_{t+1},Z_{t+1}\right)-\bar{V}^{aj}\left(L_{t+1},Z_{t+1}\right)\right]-\zeta^{aj,n}$$

are the present-value changes in the continuation value net of the moving costs.

Taking the average of (1) across households yields the expected lifetime value of a representative household in each group by age and job of

$$V^{aj}\left(L_t,Z_t\right) = \omega^{aj}\left(L_t,Z_t\right)+\beta\mathbb{E}_t\left[\bar{V}^{aj}\left(L_{t+1},Z_{t+1}\right)\right]+\nu\Omega(\bar{\varepsilon}_t^{aj}), \tag{31}$$

with the respective specializations for the age groups at the boundaries of the spectrum. Iterating forward, taking the first difference with respect to the value of being in another job, substituting for $\nu\bar{\varepsilon}_{t+1}^{aj,n}+\zeta_{t+1}^{aj,n}=\beta\mathbb{E}_{t+1}\left[\bar{V}^{an}\left(L_{t+2},Z_{t+2}\right)-\bar{V}^{aj}\left(L_{t+2},Z_{t+2}\right)\right]$ and taking the expectation at time t, hence, substituting for $\nu\bar{\varepsilon}_t^{aj,n}+\zeta_t^{aj,n}=\beta\mathbb{E}_t\left[\bar{V}^{an}\left(L_{t+1},Z_{t+1}\right)-\bar{V}^{aj}\left(L_{t+1},Z_{t+1}\right)\right]$, yields the Euler equation:

$$\begin{aligned}&\nu\bar{\varepsilon}_t^{aj,n}+\zeta_t^{aj,n}= \\ &\beta\mathbb{E}_t\left\{\left[\omega^{an}\left(L_{t+1},Z_{t+1}\right)-\omega^{aj}\left(L_{t+1},Z_{t+1}\right)\right]+\left[\nu\bar{\varepsilon}_{t+1}^{aj,n}+\zeta^{aj,n}\right]+\left[\nu\Omega(\bar{\varepsilon}_{t+1}^{an})-\nu\Omega(\bar{\varepsilon}_{t+1}^{aj})\right]\right\}.\end{aligned} \tag{32}$$

Equation (32) is an intertemporal optimality condition which says that for the marginal mover from job j to n the total cost of moving is equal to the discounted expected benefit from being in n instead of j in the following period.

A.1 Parametrization of the distribution $F(\varepsilon)$

For the sake of tractability, assume that $\varepsilon\sim Gumbel(-\gamma,1)$ where $\gamma\cong 0.5772$ is the Euler-Mascheroni constant. The first moment is equal to zero and the variance is equal to $\pi^2/6$. The c.d.f. and p.d.f. are given by $F(\varepsilon)=e^{-e^{-(\varepsilon+\gamma)}}$ and $f(\varepsilon)=e^{-(\varepsilon+\gamma)}F(\varepsilon)$. Define the auxiliary variables $\delta_t^{aj,ik}=\bar{\varepsilon}_t^{aj,i}-\bar{\varepsilon}_t^{aj,k}$, $\lambda_t^{aj,i}=\ln\left(\sum_{k\in J}e^{-\delta_t^{aj,ik}}\right)$. The probability that, when moving from a certain job j, a given job i is preferred to any other job k is given by

$$\int_{-\infty}^{\infty}f(\varepsilon)\prod_{k\neq i}F(\varepsilon+\delta_t^{aj,ik})d\varepsilon = e^{-\lambda_t^{aj,i}}=m_t^{aj,i}, \tag{33}$$

where we proceeded with a change of variable $y=\varepsilon-\lambda_t^{aj,i}$, and $\int_{-\infty}^{\infty}e^{-(y+\gamma)}e^{-\left[e^{-(y+\gamma)}\right]}dy=1$. By the law of large numbers, $m_t^{aj,i}\in(0,1)$ is the fraction of households in job j who move to job i at the end of period t. The option value is given by

$$\begin{aligned}\Omega(\bar{\varepsilon}_t^{aj}) &= \sum_{i\in J}\int_{-\infty}^{\infty}(\varepsilon+\bar{\varepsilon}_t^{aj,i})f(\varepsilon)\prod_{m\neq i}F(\varepsilon+\delta_t^{aj,im})d\varepsilon \\ &= \lambda_t^{aj,j}=-\ln\left(m_t^{aj,j}\right),\end{aligned} \tag{34}$$

where $\int_{-\infty}^{\infty}(y+\gamma)e^{-(y+\gamma)}e^{-\left[e^{-(y+\gamma)}\right]}dy = \gamma$ and $\int_{-\infty}^{\infty}(\lambda_t^{aj,j}-\gamma)e^{-(y+\gamma)}e^{-\left[e^{-(y+\gamma)}\right]}dy = \lambda_t^{aj,j} - \gamma$, and we have substituted back for $\lambda_t^{aj,j} = \ln\left(\sum_{k\in J} e^{-\delta_t^{aj,jk}}\right)$. Substituting for $\lambda_t^{aj,i} = \ln\left(\sum_{k\in J} e^{-\delta_t^{aj,ik}}\right)$ and then $\delta_t^{aj,ik} = \bar{\varepsilon}_t^{aj,i} - \bar{\varepsilon}_t^{aj,k}$ in the expression for the fraction of movers yields the probability that, when moving from job j, a certain job i is preferred to any other job k,

$$m_t^{aj,i} = \frac{e^{\bar{\varepsilon}_t^{aj,i}}}{\sum_{k\in J} e^{\bar{\varepsilon}_t^{aj,k}}}, \tag{35}$$

and the option value of moving starting from a certain job j

$$\Omega(\bar{\varepsilon}_t^{aj}) = -\ln\left(m_t^{aj,j}\right) \tag{36}$$

for $a = \{b, y, o\}$. By the law of large numbers, $m_t^{aj,i} \in (0,1)$ is the fraction of households in job j, who move to job i at the end of period t.

A.2 Rewriting the value functions

Substituting for $\nu\Omega(\bar{\varepsilon}_t^{aj}) = \nu\lambda_t^{aj,j}$ in the Bellman equation (31), given the definition $\lambda_t^{aj,j} = \ln\left(\sum_{k\in J} e^{\bar{\varepsilon}_t^{aj,k}}\right)$, the equilibrium expression for $\bar{\varepsilon}_t^{aj,k}$ in (29) and the equilibrium option value (36), yields the set of Bellman equations:

$$\begin{aligned}
V^{bj}(L_t, Z_t) &= \nu \ln\left(\sum_{k\in J} e^{\frac{1}{\nu}\left(\beta\mathbb{E}_t\left[V^{yk}(L_{t+1},s_{t+1})\right]-\zeta_t^{bj,k}\right)}\right), \\
V^{yj}(L_t, Z_t) &= \omega^j(L_t, Z_t) + \nu \ln\left(\sum_{k\in J} e^{\frac{1}{\nu}\left(\beta\mathbb{E}_t\left[(1-\lambda^y)V^{yk}(L_{t+1},s_{t+1})+\lambda^y V^{ok}(L_{t+1},s_{t+1})\right]-\zeta_t^{yj,k}\right)}\right), \\
V^{oj}(L_t, Z_t) &= \omega^j(L_t, Z_t) + \nu \ln\left(\sum_{k\in J} e^{\frac{1}{\nu}\left(\beta\mathbb{E}_t\left[(1-\lambda^o)V^{ok}(L_{t+1},s_{t+1})\right]-\zeta_t^{oj,k}\right)}\right),
\end{aligned} \tag{37}$$

that are summarized in (2).

A.3 Stationarity condition on aggregate population size

The fractions of movers have to sum to unity, thus, after reallocation, the total endowment of old incumbents is $L_t^o = \sum_{j=1}^J \sum_{i=1}^J (m_t^{oi,j} L_t^{oi})$ and that of the young incumbents is $L_t^y = \sum_{j=1}^J \sum_{i=1}^J (m_t^{yi,j} L_t^{yi})$. They are equal to their respective initial measures under stationarity. The measure of the newborn is proportional to the young incumbent population $\lambda^* L_t^y$, as $L_t^y = \sum_{j=1}^J \sum_{i=1}^J (m_t^{bi,j} L_t^{yi})$. The measure of households who die at time t is $\lambda^o L_t^o$. It follows that $\sum_{j=1}^J L_{t+1}^j = \sum_{j=1}^J L_t^j + \lambda^* L_t^y - \lambda^o L_t^o$. Labor as a fraction of the population across the J

regions, $\bar{L}$, then sums up to unity, $\bar{L}\sum_{j=1}^{J} L_t^j = 1$. The birth rate ensuring a constant world population is $\lambda^* = \lambda^o L_t^o/(\bar{L} - L_t^o)$.

The aggregate measure of young households in period $t+1$ is $L_{t+1}^y = [(1-\lambda^y) + \lambda^b]L_t^y$. The aggregate measure of old households in period $t+1$ is $L_{t+1}^o = \lambda^y L_t^y + (1-\lambda^o)L_t^o$. The first-order difference equation $L_{t+1}^o - L_t^o = \lambda^y \bar{L} - (\lambda^y + \lambda^o)L_t^o$ holds and shows that the composition of aggregate population by age group is constant over time if and only if the initial fraction of old population is equal to $L_t^o/\bar{L} = \lambda^y/(\lambda^y + \lambda^o)$. This implies $\lambda^* \equiv \lambda^y$ in every period t.

B Within-period equilibrium

In this section, we outline the production side of the economy. As in Caliendo and Parro (2015), firms in each sector and region produce intermediate goods using labor, structures, and materials. Intermediate goods are purchased by local bundlers in all regions, transforming them into a composite sectoral good that is non-tradable and sold locally to producers as an input and to consumers as a final good.

B.1 Production

Firms in each sector and country produce varieties of an intermediate good, hiring labor in different occupations $k = 1, ..., K$, employing inputs from all sectors $s = 1, ..., S$, and using structures. Apart from bundled intermediate goods, firms in region r employ two local factors, labor and structures. Labor is mobile between regions and sectors, structures are in fixed supply in each region and sector, and the intermediate goods are purchased by local bundlers at a price that includes transportation costs. Local goods bundlers do not charge any extra fees and do not have costs.

Total factor productivity in the production of intermediates is a composite of a sector-region *fundamental total factor productivity* A_t^{rs} and a variety-specific *efficiency* z^{rs}. We index varieties by their efficiency, as is customary. The output of variety z^{rs}, $q_t^{rs}(z^{rs})$, is determined as

$$q_t^{rs}(z^{rs}) = z^{rs} A_t^{rs} \left[(h_t^{rs})^{\xi^{rs}} \left(\prod_{k=1}^{K} \left(l_t^{rsk} \right)^{\epsilon_{sk}} \right)^{1-\xi^{rs}} \right]^{\gamma^{rs}} \prod_{s'=1}^{S} \left(M_t^{rs,rs'} \right)^{\gamma^{rs,rs'}}, \tag{38}$$

where l_t^{rsk} and h_t^{rs} denote labor input of occupation k and structures, respectively, and $M_t^{rs,rs'}$ denotes the inputs from sector s' available in region r and used in sector s. The parameter $\xi^{rs} \in (0,1)$ is the share of structures in value added of region-sector rs. The coefficient $\epsilon_{sk} \in (0,1)$ measures the share of labor in occupation k and sector s in all labor costs, $\sum_{k=1}^{K} \epsilon_{sk} = 1$. The coefficient γ^{rs} denotes the share of value added in region-sector rs. The coefficient $\gamma^{rs,rs'} \in (0,1)$ is the cost share of material inputs sourced in sector s' in all material input costs in the production in rs. In order to ensure constant returns to scale, we assume that $\sum_{s'=1}^{S} \gamma^{rs,rs'} = 1 - \gamma^{rs}$.

Let w_t^{rsk} be the wage earned by a worker of occupation k in region-sector rs, call ρ_t^{rs} the rental price of structures in rs, and let $P_t^{rs'}$ be the price of material inputs in r sourced from sector s'. A firm producing in region r and sector s employs a factor bundle, whose unit price is

$$x_t^{rs} = B^{rs}\left[\left(\rho_t^{rs}\right)^{\xi^{rs}}\left(\prod_{k=1}^{K}\left(w_t^{rsk}\right)^{\epsilon_{sk}}\right)^{1-\xi^{rs}}\right]^{\gamma^{rs}}\prod_{s'=1}^{S}\left(P_t^{rs'}\right)^{\gamma^{rs,rs'}}, \tag{39}$$

where $B^{rs} = \left(\xi^{rs}\gamma^{rs}\right)^{-\xi^{rs}\gamma^{rs}}\prod_{k=1}^{K}\left(\epsilon_{sk}(1-\xi^{rs})\gamma^{rs}\right)^{-\epsilon_{sk}(1-\xi^{rs})\gamma^{rs}}\prod_{s'=1}^{S}\left(\gamma^{rs,rs'}\right)^{-\gamma^{rs,rs'}}$ is a constant. The total factor productivity of a firm z in region-sector rs is given by the product $z^{rs}A_t^{rs}$. The marginal (and average) cost to produce variety z^{rs} is given by $x_t^{rs}/[z^{rs}A_t^{rs}]$.

Intermediate goods that are produced in a certain region can be traded across regions at a cost. Shipping goods from the region where they are produced, say r', to another region where they are used, say r, involves an iceberg-type trade cost, $\tau_t^{rs,r's} \geq 1$.[31] The marginal cost of variety $z^{r's}$ produced in region-sector sector $r's$ gross of the shipping costs to destination region r is given by $\left(\frac{\tau_t^{rs,r's}x_t^{r's}}{A_t^{r's}z^{r's}}\right)$. The procurement of inputs is perfectly competitive, such that in every destination region r producers source from the cheapest input supplier in the economy, gross of trade costs. The price of a specific variety of intermediate good z^s of sector s in region r is given by the lowest price at which the good is available in the region:

$$p_t^{rs}\left(z^s\right) = \min_{r'}\left\{\frac{\tau_t^{rs,r's}x_t^{r's}}{z^{r's}A_t^{r's}}\right\}. \tag{40}$$

The distribution of productivity z^{rs} in a region and sector follows a Fréchet distribution with cumulative density function $\mathcal{F}\left(z^{rs}\right) = \exp\left(-\left(z^{rs}\right)^{-\theta^s}\right)$. The joint distribution across regions is given by $\mathcal{F}\left(z^s\right) = \exp\left(-\sum_{r=1}^{R}\left(z^{rs}\right)^{-\theta^s}\right)$. The implied distribution of prices in region r for varieties of intermediate goods used in sector s at time t is given by:[32]

$$G_t^{rs}(p) = 1-\exp\left(-\Phi_t^{rs}p^{\theta_s}\right), \qquad \phi_t^{rs,r's} = \left(\frac{A_t^{r's}}{\tau_t^{rs,r's}x_t^{r's}}\right)^{\theta_s}, \qquad \Phi_t^{rs} = \sum_{r'=1}^{R}\phi_t^{rs,r's}. \tag{41}$$

[31] We follow the notation in Caliendo et al. (2019), the superscript of trade costs is labeled as {destination, source}.

[32] The distribution of the price at which region r' serves region r is given by $G^{r's,rs}(h) = 1-\mathcal{F}\left(\frac{\tau_t^{rs,r's}x_t^{r's}}{A_t^{r's}h}\right)$. The distribution of the minimum price is $\Pr\left(\min\{...,p^{r's,rs},...\}\leq p\right)$, thus it is the distribution of the event that at least a price value out of $r'=1,...,R$ is lower than p. This is the complement at 1 of the probability that all price values are greater than p, that is $\Pr\left(\min\{...,p^{r's,rs},...\}\leq p\right) = 1-\prod_{r=1}^{R}\left[1-G^{r's,rs}(p)\right] = 1-\prod_{r=1}^{R}\mathcal{F}\left(\frac{\tau_t^{rs,r's}x_t^{r's}}{A_t^{r's}p}\right)$.

We call $\tilde{q}_t^{rs}(z^s)$ the quantity of an intermediate good that is a sector-s input in region r and sourced from the supplier of the cheapest variety at the price $p_t^{rs}(z^{rs})$ in (40). The distribution of prices for intermediate goods (41) takes into account all varieties of goods of that sector s that are available in region r, whether locally produced or traded.

The varieties from sector s around the world are not available directly to producers and consumers in a specific region r, but they are made available by the local bundlers. These bundlers aggregate the varieties and then sell them on locally as intermediates to producers or as final goods to consumers. Use Q_t^{rs} to refer to the composite of sector-s goods available in region r at t through the bundler. Using a CES technology with elasticity of substitution $\eta^{rs} > 1$ in the aggregation of varieties, we obtain

$$Q_t^{rs} = \left[\int \left(\tilde{q}_t^{rs}(z^s) \right)^{1-1/\eta^{rs}} d\mathcal{F}(z^s) \right]^{\eta^{rs}/(\eta^{rs}-1)} . \tag{42}$$

The CES structure in the variety aggregation and the Fréchet distribution of productivity together yield the price index of the composite good of sector s in region r as

$$P_t^{rs} = \Gamma^{rs} \left(\sum_{r'=1}^{R} \left(x_t^{r's} \tau_t^{rs,r's} \right)^{-\theta^s} \left(A_t^{r's} \right)^{\theta^s} \right)^{-\frac{1}{\theta^s}} = \Gamma^{rs} \left(\Phi_t^{rs} \right)^{-\frac{1}{\theta^s}} , \tag{43}$$

where $\Gamma^{rs} = \Gamma\left(1 + (1 - \eta^{rs}/\theta^s)\right)$ and $\Gamma(\cdot)$ is the Gamma function and we assume $1+\theta^s > \eta^{rs}$. As mentioned above, Q_t^{rs} has two uses: it is consumed by households as a final good and it is employed by producers as an input, both being located in region r. P_t^{rs} is the sector-specific ideal consumer-price index of households in r, and it is the price of material inputs sourced from sector s and region r.

Households allocate consumption across sectoral goods according to Cobb-Douglas preferences, such that they spend a fixed share of their income, $\alpha^s > 0$ with $\sum_{s=1}^{S} \alpha_s = 1$, on the consumption of goods from sector s. The aggregate ideal consumer-price index of households in region r is then

$$P_t^r = \prod_{s=1}^{S} \left(\frac{P_t^{rs}}{\alpha_s} \right)^{\alpha_s} . \tag{44}$$

B.2 Optimal sourcing policy

Call $X_t^{rs} = P_t^{rs} Q_t^{rs}$ the sales of sectoral good s in region r. Since the output market is competitive, sales are equal to costs. Sectoral good producers combine intermediate inputs (they do not directly employ any labor), thus, total sales are equal to the total value of purchased intermediate inputs from all sectors and regions.

The probability that producers of intermediates from the source region-sector $r's$ are cheapest in the destination region-sector rs yields the cost share of producers in region-

sector rs pertaining to such intermediates:

$$\pi_t^{rs,r's} = \frac{\left(x_t^{r's}\tau_t^{rs,r's}\right)^{-\theta^s}\left(A_t^{r's}\right)^{\theta^s}}{\sum_{r''=1}^{R}\left(x_t^{r''s}\tau_t^{rs,r''s}\right)^{-\theta^s}\left(A_t^{r''s}\right)^{\theta^s}} = \frac{\phi_t^{rs,r's}}{\Phi_t^{rs}}. \tag{45}$$

The equilibrium condition (45) illustrates within-sector bilateral trade patterns between regions. Since only intermediate goods are traded, while composite sectoral goods are not, this equation fully characterizes the composition and direction of trade flows.

B.3 Output-market clearing

Use $X_t^{r's'}$ for the total value of the locally bundled good in region-sector $r's'$. Then, the expenditure of region r' on intermediates of sector s' produced in region r is $\pi_t^{r's',rs'}X_t^{r's'}$. Summing over all destinations, $\sum_{r'}^{R}\pi_t^{r's',rs'}X_t^{r's'}$ yields the total value of output (and, hence, total production costs) of intermediates produced in the sourcing region-sector rs'.

The coefficient $\gamma^{rs',rs}$ measures the fraction of total cost in region-sector rs' accruing to inputs $M_t^{rs',rs}$ sourced from the local bundler for sector s in region r; see (38). Thus, $\gamma^{rs',rs}\left(\sum_{r'}^{R}\pi_t^{r's',rs'}X_t^{r's'}\right)$ measures the purchase value of intermediates sourced from the bundler for sector s in region r and used as an input for the production of intermediates in region-sector rs'. Summing over all sectors s' in r that purchase inputs from that bundler s in r yields the aggregate sales from the bundler s in r as intermediates to local producers, $\sum_{s'}^{S}\gamma^{rs',rs}\left(\sum_{r'}^{R}\pi_t^{r's',rs'}X_t^{r's'}\right)$.

The remaining source of sales of the bundler of sector-s goods in regions r is due to consumption by workers and landlords (who own structures). Let us assume a common expenditure share of $\alpha^{s'}$ is spent on consumption of goods offered by the local bundler of sector-s goods in r. Workers in region-sector rs at time t, $L_t^{rs} \equiv \sum_k^K L_t^{rsk}$, earn an average wage rate $\bar{w}_t^{rs} = \left(\sum_k^K w_t^{rsk}L_t^{rsk}\right)/L_t^{rs}$, and total labor income amounts to $\bar{w}_t^{rs}L_t^{rs}$. There is a fixed endowment H^{rs} of structures in region-sector rs and a unitary mass of landlords, who cannot relocate to other regions. They own the local structures and rent them to local firms earning rents ρ_t^{rs}. Finally, we assume aggregate region-specific trade deficits, D_t^r, that are exogenous to the model as in Caliendo and Parro (2015). Region-specific trade deficits are the sum of sector-specific deficits, $D_t^r = \sum_{s=1}^{S} D_t^{rs}$, which will be endogenously determined in equilibrium, $D_t^{rs} = \sum_{r'=1}^{R}\pi_t^{rs,r's}X_t^{rs} - \sum_{r'=1}^{R}\pi_t^{r's,rs}X_t^{r's}$. Since all sales of the local bundler, X_t^{rs},are going to either producers as intermediates or to households as final

goods, the following accounting identity (which also serves as output-market clearing) holds:

$$X_t^{rs} = \underbrace{\sum_{s'=1}^{S} \gamma^{rs',rs} \left(\sum_{r'=1}^{R} \pi_t^{r's',rs'} X_t^{r's'} \right)}_{\text{purchases as input}} + \underbrace{\alpha^s \left(\sum_{s'=1}^{S} \left(\bar{w}_t^{rs'} L_t^{rs'} + \rho_t^{rs'} H^{rs'} \right) + D_t^r \right)}_{\text{purchases as consumer good}} . \tag{46}$$

B.4 Labor-market clearing

Local labor markets are spatially separated and clustered by type of occupation. Within each local labor market there is perfect competition. Therefore, total wages are equal to the component of value added by the corresponding labor types. Consider an arbitrary region r and sector s'. A fraction $\epsilon_{s'k}\gamma^{rs'}$ of the total production cost is spent on k-type labor in sector s'. The total labor cost associated with this share is $w_t^{rs'k} L_t^{rs'k}$. Summing over occupations yields the value added in region-sector rs' accruing to labor, that is $\bar{w}_t^{rs'} L_t^{rs'} = \sum_k^K w_t^{rs'k} L_t^{rs'k}$ where $L_t^{rs'} = \sum_k^K L_t^{rs'k}$. The labor-market clearing condition at the level of region-sector pairs is given by:

$$\bar{w}_t^{rs'} L_t^{rs'} = (1 - \xi^{rs'}) \gamma^{rs'} \left(\sum_{r'}^{R} \pi_t^{r's',rs'} X_t^{r's'} \right) . \tag{47}$$

B.5 Market-clearing for structures

The value of structures corresponds to the fraction of value added that is not attributed to labor:

$$\rho_t^{rs'} H^{rs'} = \xi^{rs'} \gamma^{rs'} \left(\sum_{r'}^{R} \pi_t^{r's',rs'} X_t^{r's'} \right) = \frac{\xi^{rs'}}{1 - \xi^{rs'}} \bar{w}_t^{rs'} L_t^{rs'} . \tag{48}$$

Thus, the rental price can be expressed in terms of wages and factor endowments.

C Contraction mapping

Let us introduce two definitions and two results.

Definition. A metric space $\{X, d\}$ is an ordered pair of a non-empty set X and a function $d : X \times X \to \Re_+$ that satisfies the properties of a distance on X: (i) $d(x, x) = 0$ for every $x \in X$; (ii) $d(x, y) = d(y, x)$ for every $x, y \in X$; (iii) for every $x, y \in X$ if $x \neq y$ then

$d(x,y) > 0$; (iv) for every $x, y, z \in X$, $d(x,z) + d(z,y) \geq d(x,y)$.

Definition. Let $\{X, d\}$ be a metric space. A mapping $F : X \to X$ is a contraction if there exists a constant $k \in [0,1)$ such that $d(F(x), F(y)) \leq kd(x,y)$ for all $x, y \in X$. That is, F is Lipschitz continuous for a Lipschitz constant strictly smaller than 1.

Definition. Let $\{X, d\}$ be a metric space. Let $\langle x_k \rangle$ be a sequence on X, that cosists of elements $x \in X$ indexed by natural numbers such as k. Then $\langle x_k \rangle$ is a Cauchy sequence if and only if $\forall \varepsilon > 0$ there exists a natural number N such that $d(x_n, x_m) < \varepsilon$ for all natural numbers $n, m \geq N$.

Banach Contraction-mapping Theorem. If $F : X \to X$ is a contraction on a complete metric space $\{X, d\}$, then there is exactly one solution $x^\star \in X$ such that $x^\star = F(x^\star)$. Furthermore, the solution can be obtained as the limit of the sequence $\langle x_i \rangle$ such that $x_{i+1} = F(x_i)$ for an arbitrary $x_0 \in X$ for $i \to \infty$.

The MFG (14)-(16) defines a vector $\boldsymbol{G}(\boldsymbol{X})$ of Lipschitz continuous and bounded functions on the nonempty and compact subset Ω of the real space. Call $d : \Omega \times \Omega \to \Re_+$ a distance function on Ω, such that $\{\Omega, d\}$ is a complete metric space. To prove that $\boldsymbol{G}(\boldsymbol{X})$ is a contraction on $\{\Omega, d\}$ we start by examining the behavior of the vector of Bellman equations.

Step 1. The dependence of $B_{aj}^{pq}(\boldsymbol{X})$ on $\boldsymbol{X}$ is mediated by the contingent payoffs. Thus, we examine the discrete difference in $B_{aj}^{pq}(\boldsymbol{X})$ in comparison to the discrete difference in $\varphi_{aj,n}^{pq}(\boldsymbol{X})$ for $n = 1, ..., J$ with respect to two arbitrary allocations $\boldsymbol{X}, \boldsymbol{Y} \in \Omega$

$$B_{aj}^{pq}(\boldsymbol{Y}) - B_{aj}^{pq}(\boldsymbol{X}) = \nu \left(\ln \left(\sum_{n=1}^{J} e^{\varphi_{aj,n}^{pq}(\boldsymbol{Y})} \right) - \ln \left(\sum_{n=1}^{J} e^{\varphi_{aj,n}^{pq}(\boldsymbol{X})} \right) \right).$$

Note that B_{aj}^{pq} is continuous and differentiable with respect to the contingent payoffs $\{\varphi_{aj,n}^{pq}\}_{n=1}^{J}$, and the latter are real-valued and compact-valued functions in $\boldsymbol{X}, \boldsymbol{Y} \in \Omega$. Thus, the Mean Value Theorem implies that on the segment between $\boldsymbol{X} \in \Omega$ and $\boldsymbol{Y} \in \Omega$ there exists a point $\boldsymbol{\rho} = (1 - \psi)\boldsymbol{X} + \psi \boldsymbol{Y}$ for some $\psi \in (0,1)$, such that

$$\begin{aligned} \left| B_{aj}^{pq}(\boldsymbol{Y}) - B_{aj}^{pq}(\boldsymbol{X}) \right| &\leq \nu \left| \sum_{n=1}^{J} \left(\frac{e^{\varphi_{aj,n}^{pq}(\boldsymbol{\rho})}}{\sum_{m=1}^{J} e^{\varphi_{aj,m}^{pq}(\boldsymbol{\rho})}} \right) \left(\varphi_{aj,n}^{pq}(\boldsymbol{Y}) - \varphi_{aj,n}^{pq}(\boldsymbol{X}) \right) \right| \\ &\leq \nu \sum_{n=1}^{J} \left(\frac{e^{\varphi_{aj,n}^{pq}(\boldsymbol{\rho})}}{\sum_{m=1}^{J} e^{\varphi_{aj,m}^{pq}(\boldsymbol{\rho})}} \right) \left| \varphi_{aj,n}^{pq}(\boldsymbol{Y}) - \varphi_{aj,n}^{pq}(\boldsymbol{X}) \right| \\ &\leq \nu \max_{n} \left| \varphi_{aj,n}^{pq}(\boldsymbol{Y}) - \varphi_{aj,n}^{pq}(\boldsymbol{X}) \right|, \quad \forall \, (a,j,p,q) \end{aligned} \tag{49}$$

where the second line is implied by the triangle inequality and the third line is implied by

the weighted average with ratios $\left(e^{\varphi^{pq}_{aj,n}(\boldsymbol{\rho})}/\sum_{m=1}^{J} e^{\varphi^{pq}_{aj,m}(\boldsymbol{\rho})}\right) \in (0,1)$ for every $n = 1, ..., J$.

Step 2. To examine the discrete difference of payoffs (13), indicate the partitions in $\boldsymbol{X}$ of frequency distribution and values, respectively, as $\mathcal{L}^{pq} \equiv X^{pq}(\mathcal{L})$ and $\mathcal{V}^{p'q'}_{an} \equiv X^{p'q'}_{an}(\mathcal{V})$, such that

$$\begin{aligned}\varphi^{pq}_{aj,n}(\boldsymbol{X}) &= \frac{1}{\nu}\sum_{p'=1}^{P} b_{p'}(X^{pq}(\mathcal{L}))\sum_{q'=1}^{Q}\pi_{qq'}\left(\beta\left[(1-\lambda^a)\,X^{p'q'}_{an}(\mathcal{V}) + \lambda^a X^{p'q'}_{(a+1)n}(\mathcal{V})\right] - \zeta^{aj,n}\right)\\ &= \frac{1}{\nu}\sum_{q'=1}^{Q}\pi_{qq'}\sum_{a'=1}^{a+1}\phi_{aa'}\sum_{p'=1}^{P} b_{p'}(X^{pq}(\mathcal{L}))\left(\beta X^{p'q'}_{a'n}(\mathcal{V}) - \zeta^{aj,n}\right),\end{aligned}$$

where in the second line we have used $\phi_{aa'} = \{1-\lambda^a, \lambda^a\}$ respectively for $a' = \{a, a+1\}$. Thus, the discrete difference in payoffs with respect to two points $\boldsymbol{X}, \boldsymbol{Y} \in \Omega$ is given by:

$$\begin{aligned}&\varphi^{pq}_{aj,n}(\boldsymbol{X}) - \varphi^{pq}_{aj,n}(\boldsymbol{Y})\\ &= \frac{\beta}{\nu}\sum_{q'=1}^{Q}\pi_{qq'}\sum_{a'=a}^{a+1}\phi_{aa'}\sum_{p'=1}^{P}\left(b_{p'}(X^{pq}(\mathcal{L}))X^{p'q'}_{a'n}(\mathcal{V}) - b_{p'}(Y^{pq}(\mathcal{L}))Y^{p'q'}_{a'n}(\mathcal{V})\right)\\ &= \frac{\beta}{\nu}\sum_{q'=1}^{Q}\pi_{qq'}\sum_{a'=a}^{a+1}\phi_{aa'}\left(\sum_{p'_x=1}^{P} b_{p'_x}(X^{pq}(\mathcal{L}))X^{p'_xq'}_{a'n}(\mathcal{V}) - \sum_{p'_y=1}^{P} b_{p'_y}(Y^{pq}(\mathcal{L}))Y^{p'_yq'}_{a'n}(\mathcal{V})\right),\\ &= \frac{\beta}{\nu}\left(\sum_{q'=1}^{Q}\sum_{a'=a}^{a+1}\sum_{p'_x=1}^{P}\pi_{qq'}\phi_{aa'}b_{p'_x}(X^{pq}(\mathcal{L}))X^{p'_xq'}_{a'n}(\mathcal{V}) - \sum_{q'=1}^{Q}\sum_{a'=a}^{a+1}\sum_{p'_y=1}^{P}\pi_{qq'}\phi_{aa'}b_{p'_y}(Y^{pq}(\mathcal{L}))Y^{p'_yq'}_{a'n}(\mathcal{V})\right),\end{aligned}$$

where in the second line we emphasize that the summation on $p' = 1, ..., P$ makes use of barycentric weights to compute the difference between averages of $X^{p'q'}_{a'n}(\mathcal{V})$ and $Y^{p'q'}_{a'n}(\mathcal{V})$ over $p' = 1, ..., P$ for a given pair (q', a'). Hence, as shown in the last line, the difference in payoffs is equal to the difference between averages of $X^{p'q'}_{a'n}(\mathcal{V})$ over (q', a', p'_x) and $Y^{p'q'}_{a'n}(\mathcal{V})$ over (q', a', p'_y). Taking the Euclidean distance, shows that the distance between the payoffs computed on points $\boldsymbol{X}$ and $\boldsymbol{Y}$, contingent on a given state (a, j, p, q) and for a given action $\bar{n} = 1, ..., J$, is equal to the distance between the corresponding means over (q', a', p') and the latter is bounded above by the maximum distance between any two elements $X^{p'_xq'}_{a'n}(\mathcal{V})$

and $Y_{a'n}^{p'_y q'}(\mathcal{V})$

$$\begin{aligned}
&\left|\varphi_{aj,n}^{pq}(\boldsymbol{X}) - \varphi_{aj,n}^{pq}(\boldsymbol{Y})\right| \qquad (50)\\
&= \frac{\beta}{\nu}\left|\sum_{q'=1}^{Q}\sum_{a'=a}^{a+1}\sum_{p'_x=1}^{P}\pi_{qq'}\phi_{aa'}b_{p'_x}(X^{pq}(\mathcal{L}))X_{a'n}^{p'_x q'}(\mathcal{V}) - \sum_{q'=1}^{Q}\sum_{a'=a}^{a+1}\sum_{p'_y=1}^{P}\pi_{qq'}\phi_{aa'}b_{p'_y}(Y^{pq}(\mathcal{L}))Y_{a'n}^{p'_y q'}(\mathcal{V})\right|\\
&\leq \frac{\beta}{\nu}\max_{(q',a',p'_x,p'_y)}\left|X_{a'n}^{p'_x q'}(\mathcal{V}) - Y_{a'n}^{p'_y q'}(\mathcal{V})\right| \ , \quad \forall\ (a,j,p,q,n).
\end{aligned}$$

Step 3. Combining (49) and (50) shows that for every grid-point (pq) and individual state (aj) the Euclidean distance between each entry of the vector image of the system of Bellman equations is bounded above by:

$$\left|B_{aj}^{pq}(\boldsymbol{Y}) - B_{aj}^{pq}(\boldsymbol{X})\right| \leq \beta \max_{n}\left\{\max_{(q',a',p'_x,p'_y)}\left|X_{a'n}^{p'_x q'}(\mathcal{V}) - Y_{a'n}^{p'_y q'}(\mathcal{V})\right|\right\} \ , \quad \forall\ (a,j,p,q).$$

Sort $\boldsymbol{B}(\boldsymbol{X}) = \{B_{aj}^{pq}(\boldsymbol{X}), \forall(a,j,p,q)\}$, $\boldsymbol{B}(\boldsymbol{Y}) = \{B_{aj}^{pq}(\boldsymbol{Y}), \forall(a,j,p,q)\}$, $\boldsymbol{X}(\mathcal{V})$ and $\boldsymbol{Y}(\mathcal{V})$ as matrices of dimension $[J \times 2 \cdot Q \cdot P]$, with jobs on the rows and occurrences of the triplet (q,a,p) on the columns. Define the function $d_B : \Re_+^{[J\times 2\cdot Q\cdot P]} \times \Re_+^{[J\times 2\cdot Q\cdot P]} \to \Re_+$

$$d_B(Y,X) = \max_{r}\max_{c_x,c_y}|Y_{rc_y} - X_{rc_x}| \ , \quad \forall X, Y: \begin{array}{l} X = \{X_{rc} \in \Re_+ : r = 1,...,J \text{ and } c = 1,...,2\cdot Q\cdot P\} \\ Y = \{Y_{rc} \in \Re_+ : r = 1,...,J \text{ and } c = 1,...,2\cdot Q\cdot P\} \end{array}$$

that satisfies the properties of a distance: $d_B(Y,X) \geq 0$, $d_B(Y,X) = 0 \iff X = Y$, $d_B(Y,X) = d_B(X,Y)$ and $d_B(Y,Z) + d_B(Z,X) \leq d(Y,X)$. Furthermore, $d_B(kY,kX) = |k|d_B(Y,X)$ for every arbitrary $k \in \Re$, thus, d_B is a norm on the vector space $\Re_+^{[J\times 2\cdot Q\cdot P]}$. The system of (49) and (50) for all occurrences (a,j,p,q) implies

$$d_B\left(\boldsymbol{B}(\boldsymbol{Y}), \boldsymbol{B}(\boldsymbol{X})\right) \leq \beta d_B\left(\boldsymbol{Y}(\mathcal{V}), \boldsymbol{X}(\mathcal{V})\right) \ , \quad \forall\ \boldsymbol{X}(\mathcal{V}), \boldsymbol{Y}(\mathcal{V}) \in \Re_+^{[J\times 2\cdot Q\cdot P]}. \qquad (51)$$

Therefore, for every $\beta \in [0,1)$ the system of Bellman equations is a contraction on the complete metric space $\{\Re_+^{[J\times 2\cdot Q\cdot P]}, d_B\}$.

Step 4. Sort $\boldsymbol{K}(\boldsymbol{X}) = \{K_{aj}^{pq}(\boldsymbol{X}), \forall(a,j,p,q)\}$ and $\boldsymbol{X}(\mathcal{L})$ as $[J \times 2 \cdot Q \cdot P]$ matrices, with jobs on the rows and occurrences of the triplet (q,a,p) on the columns. The column stacks $\boldsymbol{G}(\boldsymbol{X}) = [\boldsymbol{B}(\boldsymbol{X}); \boldsymbol{K}(\boldsymbol{X})]$ and $\boldsymbol{X} = [\boldsymbol{X}(\mathcal{V}); \boldsymbol{X}(\mathcal{L})]$ are matrices of dimension $[2 \cdot J \times 2 \cdot Q \cdot P]$, and the space Ω can be defined accordingly as a compact subset in $\Re_+^{[2\cdot J\times 2\cdot Q\cdot P]}$. This allows the iterative scheme for the MFG $\boldsymbol{X}^{i+1} = \boldsymbol{G}(\boldsymbol{X}^i)$ to be written as the vector equation:

$$\begin{bmatrix} \boldsymbol{X}^{i+1}(\mathcal{V}) \\ \boldsymbol{X}^{i+1}(\mathcal{L}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}(\boldsymbol{X}^i) \\ \boldsymbol{K}(\boldsymbol{X}^i) \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}\left([\boldsymbol{X}^i(\mathcal{V}); \boldsymbol{X}^i(\mathcal{L})]\right) \\ \boldsymbol{K}\left([\boldsymbol{X}^i(\mathcal{V}); \boldsymbol{X}^i(\mathcal{L})]\right) \end{bmatrix} \ , \quad \text{given } \boldsymbol{X}^0 \in \Omega. \qquad (52)$$

As it can be seen from (50), at each iteration the image of the system of Bellman equations, i.e. $\boldsymbol{X}^{i+1}(\mathcal{V})$, depends on $\boldsymbol{X}^{i}(\mathcal{L})$ through the barycentric weights only, which implies that the mapping $\boldsymbol{B}(\boldsymbol{X}^{i})$ contracts with respect to d_B, as shown in (51), for every vector $\boldsymbol{X}^{i}(\mathcal{L})$. Thus, the iteration on (52) generates a sequence of values of the DOCP $(\boldsymbol{X}^{1}(\mathcal{V}), \boldsymbol{X}^{2}(\mathcal{V}), ..., \boldsymbol{X}^{i}(\mathcal{V}), \boldsymbol{X}^{i+1}(\mathcal{V}), ...)$ that is Cauchy with respect to d_B, hence, its limit is unique in the complete metric space $\{\Re_{+}^{[J\times 2\cdot Q\cdot P]}, d_B\}$.

Step 5. To conclude the proof, we must extend the previous result on $\Omega \subset \Re_{+}^{[2\cdot J\times 2\cdot Q\cdot P]}$ with respect to a norm d on Ω. This will show that the mapping $\boldsymbol{X}^{i+1} = \boldsymbol{G}(\boldsymbol{X}^{i})$ is a contraction on a complete metric space $\{\Omega, d\}$. To this purpose, define a function $d : \Omega \times \Omega \to \Re_{+}$ as given by the subdistance on the subspace of the Bellman equations:

$$d(\boldsymbol{Y}, \boldsymbol{X}) = d_B(\boldsymbol{Y}(\mathcal{V}), \boldsymbol{X}(\mathcal{V})) \ , \quad \forall \ \boldsymbol{X}, \boldsymbol{Y} \in \Omega \subset \Re^{[2\cdot J\times 2\cdot Q\cdot P]}.$$

The following properties hold: $d(\boldsymbol{Y}, \boldsymbol{X}) \geq 0$, $d(\boldsymbol{Y}, \boldsymbol{X}) = d(\boldsymbol{X}, \boldsymbol{Y})$, $d(\boldsymbol{Y}, \boldsymbol{Z}) + d(\boldsymbol{Z}, \boldsymbol{X}) \leq d(\boldsymbol{Y}, \boldsymbol{X})$ and $d(k\boldsymbol{Y}, k\boldsymbol{X}) = |k| d(\boldsymbol{Y}, \boldsymbol{X})$ for every scalar k hold. The last requirement for d being a valid norm on Ω is that $d(\boldsymbol{Y}, \boldsymbol{X}) = 0 \iff \boldsymbol{Y} = \boldsymbol{X}$. It is obvious to conclude that $\boldsymbol{Y} = \boldsymbol{X} \implies d(\boldsymbol{Y}, \boldsymbol{X}) = 0$, thus, the other direction of causality is the one that should be proved. Every $\boldsymbol{Y}(\mathcal{V}) \neq \boldsymbol{X}(\mathcal{V}) \implies d(\boldsymbol{Y}, \boldsymbol{X}) > 0$ and we know already that $d_B(\boldsymbol{Y}(\mathcal{V}), \boldsymbol{X}(\mathcal{V})) = 0 \iff \boldsymbol{Y}(\mathcal{V}) = \boldsymbol{X}(\mathcal{V})$. Therefore, we look for the only possible contradiction and assume

$$\text{Hp} : \exists \ \boldsymbol{X}, \boldsymbol{Y} \in \Omega : d(\boldsymbol{Y}, \boldsymbol{X}) = 0 \text{ and } \boldsymbol{Y}(\mathcal{L}) \neq \boldsymbol{X}(\mathcal{L}).$$

From (50), taking the maximum over $n = 1, ..., J$ yields for every $\boldsymbol{X}, \boldsymbol{Y} \in \Omega$

$$\begin{aligned} \max_{n} \left| \varphi_{aj,n}^{pq}(\boldsymbol{X}) - \varphi_{aj,n}^{pq}(\boldsymbol{Y}) \right| &\leq \frac{\beta}{\nu} \max_{n} \left\{ \max_{(q',a',p'_x,p'_y)} \left| X_{a'n}^{p'_x q'}(\mathcal{V}) - Y_{a'n}^{p'_y q'}(\mathcal{V}) \right| \right\} \quad \forall \ (a, j, p, q) \\ &\leq \frac{\beta}{\nu} d_B\left(\boldsymbol{X}(\mathcal{V}), \boldsymbol{Y}(\mathcal{V})\right) \quad \forall \ (a, j, p, q). \end{aligned}$$

Therefore, $\varphi_{aj,n}^{pq}(\boldsymbol{Y}) = \varphi_{aj,n}^{pq}(\boldsymbol{X})$ for every (a, j, p, q, n) if and only if $\boldsymbol{X}(\mathcal{V}) = \boldsymbol{Y}(\mathcal{V})$. Furthermore, from (15), images of the system of Kolmogorov equations are continuous and bounded functions of the payoffs, therefore, $\varphi_{aj,n}^{pq}(\boldsymbol{Y}) = \varphi_{aj,n}^{pq}(\boldsymbol{X})$ for every (a, j, p, q, n) also implies $\boldsymbol{B}(\boldsymbol{Y}) = \boldsymbol{B}(\boldsymbol{X})$ and $\boldsymbol{K}(\boldsymbol{Y}) = \boldsymbol{K}(\boldsymbol{X})$. Let $\boldsymbol{X}(\mathcal{V}) = \boldsymbol{Y}(\mathcal{V})$, call $x_{an}^{p'q} \equiv \sum_{q'=1}^{Q} \pi_{qq'} \sum_{a'=a}^{a+1} \phi_{aa'} X_{a'n}^{p'q'}(\mathcal{V})$ and compute the difference between payoffs:

$$\varphi_{aj,n}^{pq}(\boldsymbol{X}) - \varphi_{aj,n}^{pq}(\boldsymbol{Y}) = \sum_{p'=1}^{P} x_{an}^{p'q} b_{p'}(X_{::}^{pq}(\mathcal{L})) - \sum_{p'=1}^{P} x_{an}^{p'q} b_{p'}(Y_{::}^{pq}(\mathcal{L})) \quad \forall \ (a, p, q, n).$$

Thus, interpreting barycentric weights as a histogram on the support $\{x_{an}^{pq} \ \forall (a, p, q, n)\}$, the equivalence in payoffs $\varphi_{aj,n}^{pq}(\boldsymbol{X}) = \varphi_{aj,n}^{pq}(\boldsymbol{Y})$ holds if and only if the grid of barycentric weights yields the same average under $\boldsymbol{X}$ and $\boldsymbol{Y}$, for every (a, p, q, n).

However, barycentric weights do not vary by age and action (a, n), thus, the same linear relationship should be satisfied by several age-and-action-specific occurrences such that, for every (p, q), each of the two nonzero barycentric weights $b_{k_i^x}(X^{pq}(\mathcal{L})), b_{k_i^y}(Y^{pq}(\mathcal{L}))$ for $i = 1, 2$ should be the solution to an overdetermined linear system of $2 \cdot J$ equations in 2 unknowns, where $J > 1$. Given a large set of age and action pairs, admissible solutions such that barycentric weights are equal and $\boldsymbol{X}(\mathcal{L}) \neq \boldsymbol{Y}(\mathcal{L})$ do not exist.[33] Therefore, given $\boldsymbol{X}(\mathcal{V}) = \boldsymbol{Y}(\mathcal{V})$, $\varphi^{pq}_{aj,n}(\boldsymbol{X}) = \varphi^{pq}_{aj,n}(\boldsymbol{Y}) \forall\ (a, p, q, n) \iff \boldsymbol{X}(\mathcal{L}) = \boldsymbol{Y}(\mathcal{L})$. Combining the previous results yields

$$
\begin{aligned}
d(\boldsymbol{X}, \boldsymbol{Y}) = 0 &\iff d_B(\boldsymbol{Y}(\mathcal{V}), \boldsymbol{X}(\mathcal{V})) = 0 \\
&\iff \boldsymbol{Y}(\mathcal{V}) = \boldsymbol{X}(\mathcal{V}) \\
&\iff \varphi^{pq}_{aj,n}(\boldsymbol{Y}) = \varphi^{pq}_{aj,n}(\boldsymbol{X}) \quad \forall\ (a, j, p, q, n) \\
&\implies \boldsymbol{B}(\boldsymbol{Y}) = \boldsymbol{B}(\boldsymbol{X}) \text{ and } \boldsymbol{K}(\boldsymbol{Y}) = \boldsymbol{K}(\boldsymbol{X})\ , \forall\ \boldsymbol{X}, \boldsymbol{Y} \in \Omega \\
&\implies \boldsymbol{Y}(\mathcal{L}) = \boldsymbol{X}(\mathcal{L})\ \forall\ \boldsymbol{X}, \boldsymbol{Y} \in \Omega.
\end{aligned}
$$

This contradicts Hp. Therefore, d is a norm on Ω and

$$d\left(\boldsymbol{G}(\boldsymbol{Y}), \boldsymbol{G}(\boldsymbol{X})\right) \leq \beta d\left(\boldsymbol{Y}, \boldsymbol{X}\right)$$

shows that for every $\beta \in [0, 1)$ the MFG $\boldsymbol{G}$ is a contraction on the complete metric space $\{\Omega, d\}$. Banach Contraction-mapping Theorem implies that the unique solution to (52) can be obtained iterating on $\boldsymbol{X}^{i+1} = \boldsymbol{G}(\boldsymbol{X}^i)$ for an arbitrary $\boldsymbol{X}^0 \in \Omega$. ■

D Intuition about the role of uncertainty

In this section we provide a more formal derivation of the implications of option-value approach to choices under uncertainty, discussed in Section 4.

1. *If agents are risk-averse, aggregate uncertainty depresses the incentive to relocate relative to perfect foresight.* To see this, consider an exogenous, ceteris paribus increase in the real wage paid in job n. Under risk aversion, the function V^{an} is strictly concave with respect to updates in the own real wage, and Jensen's inequality implies $\mathbb{E}_t\left[V^{an}(L_{t+1}, Z_{t+1})\right] \leq V^{an}\left(L_{t+1}, \mathbb{E}_t\left[Z_{t+1}\right]\right)$ and $\mathbb{E}_t\left[V^{(a+1)n}(L_{t+1}, Z_{t+1})\right] \leq V^{(a+1)n}\left(L_{t+1}, \mathbb{E}_t\left[Z_{t+1}\right]\right)$, which holds with equality only under perfect foresight.

2. *Due to aggregate uncertainty, more households rationally spend a greater portion of their life in relatively bad jobs.* To see this, consider two jobs $\{n, j\}$ with the same fundamentals

[33] A technical conditions on the grid of distributions $(l^1, l^2, ..., l^P)$ and barycentric weights shall be imposed, namely: *Given* $x^{p'q}_{an} \in \Re_+ : p' = 1, ..., P$ *for all* (a, q, n)*, and two states* $\boldsymbol{X}, \boldsymbol{Y} \in \Omega$*, if* $\boldsymbol{X}(\mathcal{L}) \neq \boldsymbol{Y}(\mathcal{L})$ *then there exists at least one occurrence* (a, p, q, n) *such that* $\sum_{p'=1}^{P} x^{p'q}_{an} b_{p'}(X^{pq}_{\cdot\cdot}(\mathcal{L})) \neq \sum_{p'=1}^{P} x^{p'q}_{an} b_{p'}(Y^{pq}_{\cdot\cdot}(\mathcal{L}))$. However, this is always the case if the vector of barycentric weights for $\boldsymbol{X} \in \Omega$, i.e. $\{b_{p'}(Y^{pq}(\mathcal{L})) : p' = 1, ..., P\}$ cannot be written as a linear combination of the vector of barycentric weights for $\boldsymbol{Y} \in \Omega$, i.e. $\{b_{p'}(X^{pq}(\mathcal{L})) : p' = 1, ..., P\}$. Unless the grid of distributions $(l^1, l^2, ..., l^P)$ and the distance δ used to compute barycentric weights are trivial, this is always true in large scale problems.

and identical realization of an aggregate stochastic shock at time t, such that $V^{an}(L_t, Z_t) - V^{aj}(L_t, Z_t) = 0$ and $\mathbb{E}_t\left[V^{an}(L_{t+1}, Z_{t+1}) - V^{aj}(L_{t+1}, Z_{t+1})\right] = 0$. Recall that moving decisions are made at time t, by looking at the foreseen differences in next-period values.

Consider a positive shock to the value of job n with some persistent behavior. The concavity of the value function with respect to updates in the own real wage implies that foreseen positive differences under perfect foresight are smaller than under uncertainty with rational but risk-averse agents, $V^{an}(L_{t+1}, Z_{t+1}) - V^{aj}(L_{t+1}, Z_{t+1}) > \mathbb{E}_t\left[V^{an}(L_{t+1}, Z_{t+1}) - V^{aj}(L_{t+1}, Z_{t+1})\right] > 0$. And the reverse conclusion holds for negative shocks, $V^{an}(L_{t+1}, Z_{t+1}) - V^{aj}(L_{t+1}, Z_{t+1}) < \mathbb{E}_t\left[V^{an}(L_{t+1}, Z_{t+1}) - V^{aj}(L_{t+1}, Z_{t+1})\right] < 0$.

3. *Aging reduces the option value of relocation.* Consider the value of the DOCP for an agent who moves from job j to job n within the two age spells, young and old:

$$V^{yj}(L_t, Z_t) = \omega^{yj}(L_t, Z_t) - \zeta^{yj,n} + \beta\left\{(1-\lambda^y)\mathbb{E}_t\left[V^{yn}\left(L_{t+1}, Z_{t+1}\right)\right] + \lambda^y\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right]\right\},$$
$$V^{oj}(L_t, Z_t) = \omega^{oj}(L_t, Z_t) - \zeta^{oj,n} + \beta\left\{(1-\lambda^o)\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right]\right\}.$$

Given continuous and positive functions $V^{aj} \geq 0$ for $a = \{y, o\}$ and $\omega^{yj} = \omega^{oj} \geq 0$, the sorting in moving cost $\zeta^{oj,n} \geq \zeta^{yj,n} \geq 0$ implies that $V^{yj}(L_t, Z_t) - V^{oj}(L_t, Z_t) \geq V^{yj}(L_t, Z_t) + \zeta^{yj,n} - V^{oj}(L_t, Z_t) - \zeta^{oj,n}$. Without loss of generality, we assume $\zeta^{yj,n} = \zeta^{oj,n}$. Thus, the gap between total value and continuation value does not depend on age, and

$$V^{yj}(L_t, Z_t) - V^{oj}(L_t, Z_t) = \beta\mathbb{E}_t\left[(1-\lambda^y)V^{yn}\left(L_{t+1}, Z_{t+1}\right) - (1-\lambda^o-\lambda^y)V^{on}\left(L_{t+1}, Z_{t+1}\right)\right].$$

Then, the ranking of total value and continuation value between the young and the old must be the same. Clearly, $1 - \lambda^o \leq \lambda^y$ is a sufficient condition for both total value $V^{yj}(L_t, Z_t) > V^{oj}(L_t, Z_t)$ and continuation value $(1-\lambda^y)\mathbb{E}_t\left[V^{yn}\left(L_{t+1}, Z_{t+1}\right)\right] + \lambda^y\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right] > (1-\lambda^o)\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right]$ of the young being larger.

Let this sufficient condition be violated, i.e., $\lambda^y < 1-\lambda^o$, and assume that the continuation value $(1-\lambda^y)\mathbb{E}_t\left[V^{yn}\left(L_{t+1}, Z_{t+1}\right)\right] + \lambda^y\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right] \leq (1-\lambda^o)\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right]$ is larger for the old. Since $1-\lambda^y > \lambda^o$, then, for $\mathbb{E}_t\left[V^{oj}\left(L_{t+1}, Z_{t+1}\right)\right] > \mathbb{E}_t\left[V^{yj}\left(L_{t+1}, Z_{t+1}\right)\right] > 0$, a lower bound to the continuation value of the young is given by $(1-\lambda^y)\mathbb{E}_t\left[V^{yn}\left(L_{t+1}, Z_{t+1}\right)\right] + \lambda^y\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right] > \lambda^o\mathbb{E}_t\left[V^{yn}\left(L_{t+1}, Z_{t+1}\right)\right] + (1-\lambda^o)\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right]$ and, therefore, the continuation value of the old can be larger only, if the expected next-period value for the young is negative

$$\lambda^o\mathbb{E}_t\left[V^{yn}\left(L_{t+1}, Z_{t+1}\right)\right] + (1-\lambda^o)\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right] \leq (1-\lambda^o)\mathbb{E}_t\left[V^{on}\left(L_{t+1}, Z_{t+1}\right)\right]$$
$$\implies \mathbb{E}_t\left[V^{yn}\left(L_{t+1}, Z_{t+1}\right)\right] < 0, \quad \not\exists$$

which is a contradiction. Hence, for the same job and the same moving cost between age cohorts, both the total value and the continuation value for the young are larger than for the old. The same holds true, if $\zeta^{oj,n} \geq \zeta^{yj,n}$.

E Quantification of the model

To determine the key parameters specific to the model, we need two types of ingredients. First, we need the parameters and realizations of the stochastic state determining the within-period equilibrium which will determine the counterfactual distribution of wages paid and other (input and output) prices given the distribution of individuals across jobs. The respective parameters are the production function elasticities and TFP values, trade costs, elasticities of substitution and final consumption shares. Second, we need the parameters determining the intertemporal equilibrium which will be determine the counterfactual distribution of value and individuals across jobs in each period of the dynamic problem. The respective parameters pertain to mobility costs, discounting, and the dispersion of tastes for jobs. In general, we use 2012 as a benchmark for the within-period equilibrium. In order to estimate characteristics of the TFP process, we employ panel data between the years 2003 and 2014. For estimating mobility costs from the mobility of individuals between pairs of jobs we employ data on individual workers in France and use the observed changes from jobs in 2009 to jobs in 2010 and from 2012 to 2013, respectively.

E.1 Intratemporal equilibrium parameters

The data sets employed to obtain the intratemporal parameters are the annual structural statistics of companies for France (the *FICUS-FARE* dataset published by INSEE), the French administrative employer-employee dataset *Declaration Annuelle des Donnees Sociales* published by INSEE, WIOD data on input-output tables as well as their data on socioeconomic accounts and Eurostat's European Road Freight Transport Survey. Apart from these data, we need data on bilateral sales (trade flows) between all pairs of NUTS2 regions in France and the Rest of the World (ROW) ($R = 22 + 1$) for every sector (WIOD sector $S = 49$).

Output data at this level are available from the *FICUS-FARE* dataset and from WIOD for the ROW. We measure trade flows of manufactured goods between French regions using the European Road Freight Transport Survey after mapping their product categories into WIOD's. This permits distributing output of a region and sector across other regions based on freight data. We can directly observe imports and exports with the ROW for each region and sector. To approximate services trade, we calculate distance elasticities from sector-level country-by-country WIOD data and distribute services based on the distances between French regions for each services sector, accordingly. From French administrative matched employer-employee data (DADS) and balance-sheet data (Ficus-Fare) we compute the fraction of the employment of labor in hours by region, sector, and occupation in the baseline year, that is L^{rsk}_{2012}. Using two sources of data in Eurostat, *Land use overview by NUTS 2 regions* and *Annual national accounts, Breakdowns of non-financial assets by type, industry and sector*, we measure the land use in squared kilometers by sector and regions in France. Moreover, we observe total fixed assets (in gross terms, according to ESA 2010) by sector in the country. We use the share of assets by sector times sectoral and regional land use in the year 2012 as our measure for the units of structures used by region and

sector, H^{rs}. In Panel (b) of Figure 6.1 we report the distribution of structures across French regions averaged over sectors using sector-level weights based on the employment share of each sector.

To construct the ROW analogue for all variables, we use WIOD data and scale them such that the values for France match the corresponding aggregate levels in the DADS data. This makes the value that we observe in WIOD for ROW comparable in magnitude to the data for France. For structures, where ROW data is missing, we use the relationship between production factors in France to predict them for the ROW.

Table 11: Summary statistics for ROW versus France

	mean	std.dev.	p10	p50	p90
Revenue ROW / Revenue France	22.550	2.625	19.166	22.378	26.431
Labor cost ROW / Labor cost France	63.882	17.339	46.486	58.489	92.023
Value of intermed. ROW / Value of intermed. France	22.940	3.033	19.085	22.763	26.711

Source: Own computations based on WIOD, DADS, and Eurostat.

Table 11 reports descriptive statistics of key economic aggregates for the ROW relative to France on average across years and sectors. Revenues in the ROW are about 22.5 times the ones in France; the labor force is about 64 times the one in France, and expenditures in intermediates are about 23 times the ones in France.

E.2 Trade elasticities θ^s.

The calibration requires knowledge of sector-specific trade elasticities θ^s which we take from Caliendo and Parro (2015).

E.3 Measuring trade costs $\tau^{rs,r's}$.

Trade costs are measured based on data for 2012. We calculate trade costs by the so-called Head-Ries index (Head and Ries, 2001) using interregional sales and purchases between regions r and r' in sector s, $X^{rs,r's}$ and $X^{r's,rs}$ both from year 2012 and based on the aforementioned freight data, each normalized by the intraregional absorption, $X^{rs,rs}$ and $X^{r's,'s}$:

$$\tau^{rs,r's} = \left(\frac{X^{rs,rs} X^{r's,r's}}{X^{rs,r's} X^{r's,rs}} \right)^{1/(2\theta^s)} . \tag{53}$$

We present the such obtained trade costs in Table 4, note that we ensure that $\tau^{rs,r's} \geq 1$ throughout. We will assume prohibitive trade costs equal to $\tau = 1,000$ for all cases where no trade costs are observed.

E.4 Measuring cost-share parameters ϵ^{sk}, ξ^{rs}, γ^{rs}, and $\gamma^{rs,rs'}$.

As the cost-share parameters will be treated as time-invariant, we will skip the time index in this section and only note that all data used to measure these parameters pertain to the year 2012.

We observe the total wages paid to workers from the employer-employee data set. We aggregate these data to obtain the total wages paid to persons of skill group k employed in sector s of region r, $w^{rsk}L^{rsk}$, and calculate the share paid to each skill group as $\epsilon^{sk} = (\sum_{r=1}^{R} w^{rsk}L^{rsk})/(\sum_{r=1}^{R}\sum_{k=1}^{K} w^{rsk}L^{rsk})$. Since we do not observe wages by skill groups for the ROW, we take the average of the French skill shares over regions in France for each sector of ROW. From the production data we use sector-region-level aggregates of costs paid for intermediates and raw materials (which we combine to intermediates) and salaries, respectively. Total costs that are devoted to salaries divided by total revenues will be used as a measure of $\gamma^{rs}\left(1-\xi^{rs}\right)$, and, the costs of intermediates in region r and sector s, relative to revenues correspond to $(1-\gamma^{rs})$. The production function elasticity of structures can be calculated assuming that all production function elasticities sum up to unity, $\gamma^{rs}\xi^{rs} = 1 - \gamma^{rs}\left(1-\xi^{rs}\right) - (1-\gamma^{rs})$. For the ROW, we take the sector-specific average of $\gamma^{rs}\xi^{rs}$ across all regions as an estimate for the ROW. Moreover, we calculate $(1-\gamma^{rs})$ and $\gamma^{rs}\left(1-\xi^{rs}\right)$ using the share of intermediate input purchases and labor in the total costs of production in sector s in all countries but France from WIOD data and adjust the shares such that $\gamma^{ROW,s}\xi^{ROW,s} + \left(1-\gamma^{ROW,s}\right) + \gamma^{ROW,s}\left(1-\xi^{ROW,s}\right) = 1$. Moreover, WIOD provides information on the fraction of spending of sector s on inputs from sector s' in total purchases of intermediates for each country. We use the data for France and parameterize $\gamma^{rs,rs'} = \gamma^{s,s'}$ for its regions. For the ROW, we relate the ROW purchases of s from s' to calibrate $\gamma^{rs,rs'}$ in the same way.

E.5 Measuring final consumption share α^s.

We use WIOD data on final consumption for France to obtain the shares spent on each sector, α^s, respectively. Production function elasticities and consumption shares are summarized in Table 2.

E.6 Estimation of TFP levels

Since our model suggest that the process of TFP is stochastic, we aim at obtaining a time series of the stochastic component of TFP, A_t^{rs} that allows us to obtain estimates of the long-run mean of region-sector specific productivity shocks as well as measures of uncertainty about its stochastic component. In order to obtain this time series, we use panel data for 12 years between 2003 and 2014. Once we have obtained the time series of the stochastic component and its region-sector-specific level, we specify the stochastic process of productivity shocks across French regions and sectors as an AR(1) process which exhibits an rs-specific level in the short run and calculate the long-run mean implied by the specified process for each region and sector. Finally, we measure uncertainty as the variation of the TFP residuals around the region-sector-time-specific prediction from the AR(1) process.

We proceed in three steps. First, we isolate a time series of region-sector-specific productivity. Second, we isolate the region-sector-specific time invariant level differences and the stochastic component of TFP. Finally, we estimate the stochastic AR(1) process underlying the latter time series. In measuring region-sector-specific productivity we follow Caliendo and Parro (2015) and Caliendo et al. (2018), whose production equilibrium is fully consistent with the one that we consider in this paper.[34] However, relative to Caliendo et al. (2018) we need to solve the model in levels rather than in changes, as the so-called exact-hat-algebra is not available for it. For this reason, we need to make use of the region-sector price index P_t^{rs}.

In a first step, we obtain the aggregate sector-region-level production function by relating aggregate output in a region and sector to aggregate inputs used in that sector. Aggregating over skill groups, let $l_t^{rs}(z)$, denote the demand for labor by a producer operating in region r and sector s at time t with an idiosyncratic efficiency of z. Let $M_t^{rs}(z)$ and $H_t^{rs}(z)$ denote the same producer's demand of intermediates and structures. With $p_t^{rs}(z)q_t^{rs}(z)$ denoting the sales value of that producer, and with P_t^{rs} denoting the price of the intermediates bundle in r, s, w^{rs} wages in r, s and, ρ_t^{rs} the price of structures in r, s, we obtain:

$$h_t^{rs}(z) = \gamma^{rs}\xi^{rs}\frac{p_t^{rs}(z)q_t^{rs}(z)}{\rho_t^{rs}} \qquad l_t^{rs}(z) = \gamma^{rs}\left(1-\xi^{rs}\right)\frac{p_t^{rs}(z)q_t^{rs}(z)}{w_t^{rs}} \qquad M_t^{rs}(z) = \left(1-\gamma^{rs}\right)\frac{p_t^{rs}(z)q_t^{rs}(z)}{P_t^{rs}} \ .$$

Aggregating over varieties yields

$$H_t^{rs} = \gamma^{rs}\xi^{rs}\frac{Y_t^{rs}}{\rho_t^{rs}} \qquad L_t^{rs} = \gamma^{rs}\left(1-\xi^{rs}\right)\frac{Y_t^{rs}}{w_t^{rs}} \qquad M_t^{rs} = \left(1-\gamma^{rs}\right)\frac{Y_t^{rs}}{P_t^{rs}} \ .$$

Thus, optimal factor demands can be expressed in terms of aggregate inputs, namely structures, labor and intermediate inputs in rs and t, H_t^{rs}, L_t^{rs} and M_t^{rs}, and aggregate output produced in rs and t, Y_t^{rs}, respectively:

$$h_t^{rs}(z) = \frac{p_t^{rs}(z)q_t^{rs}(z)}{Y_t^{rs}}H_t^{rs} \qquad l_t^{rs}(z) = \frac{p_t^{rs}(z)q_t^{rs}(z)}{Y_t^{rs}}L_t^{rs} \qquad M_t^{rs}(z) = \frac{p_t^{rs}(z)q_t^{rs}(z)}{Y_t^{rs}}M_t^{rs} \ .$$

Substituting in the production function of in (38) yields gross output in region r sector s, Y_t^{rs}, which scaled by the region and sector specific price index yields real gross output:

$$\frac{Y_t^{rs}}{P_t^{rs}} = \frac{x_t^{rs}}{P_t^{rs}}\left[\left(H_t^{rs}\right)^{\gamma^{rs}\xi^{rs}}\left(L_t^{rs}\right)^{\gamma^{rs}\left(1-\xi^{rs}\right)}\left(M_t^{rs}\right)^{1-\gamma^{rs}}\right]. \tag{54}$$

In the latter, we substituted for the competitive price $p_t^{rs}(z) = \frac{x_t^{rs}}{z(A_t^{rs})}$. Equation (25) shows that the model implies a region-and-sector specific production function with observed TFP given by the input-bundle price relative to the output price, x_t^{rs}/P_t^{rs}. By expressing the sectoral price index in (43) using the domestic expenditure share in (45), $\pi_t^{rs,rs}$, the observed

[34] In particular, we define trade flows and value added shares exactly as in Sections 3.3 and 4 in Caliendo and Parro (2015), and we follow Section 4 in Caliendo et al. (2018) in measuring TFP.

TFP of region r and sector s, A_t^{rs}, is determined by:

$$\frac{x_t^{rs}}{P_t^{rs}} = \frac{A_t^{rs}}{\Gamma^{rs}\left(\pi_t^{rs,rs}\right)^{\frac{1}{\theta^s}}}. \tag{55}$$

Hence, in order to obtain an estimate of A_t^{rs}, we will first estimate observed sector-region-specific TFP from a region-sector-specific aggregate production function as a residual and then rescale it using data on internal trade shares and elasticities. Specifically, we estimate $\left(\frac{x_t^{rs}}{P_t^{rs}}\right)$ using equation (25) and data on the value of gross output in region r and sector s, price indices of the composite good in region r and sector s, as well as structures, employment, material inputs and the respective output elasticities in each region and sector.

We observe data on gross output, structures, hours worked and expenditures on intermediates directly, but we need additional information on prices of output as well as the aggregate price of the intermediate good. Moreover, we will rely on the production function elasticities discussed above. In order to obtain an estimate of the output price index, we use data on trade shares across region-sectors and can recover an estimate of $\left(\frac{P_t^{rs}}{\Gamma^{rs}}\right)$ up to a constant across regions within a sector. The share of total expenditure of region r and sector s on intermediates sourced from region r' and sector s, that is $X^{rs,r's}/X_t^{rs}$, corresponds to the share $\pi_t^{rs,r's}$ defined in equation (45). Substituting for the equilibrium sectoral price index in (43) yields the fraction of spending of $\{rs\}$ on goods from $\{r's\}$:

$$X_t^{rs,r's} = \frac{\phi_t^{rs,r's}}{\left(P_t^{rs}/\Gamma^{rs}\right)^{-\theta^s}} X_t^{rs}\ . \tag{56}$$

Given (trade or) sales flows and θ^s, we use this to express the normalized import share as defined in Eaton and Kortum (2002):[35]

$$\frac{X_t^{rs,r's}/X_t^{rs}}{X_t^{r's,r's}/X_t^{r's}} \frac{1}{\left(\tau_t^{rs,r's}\right)^{-\theta^s}} = \tilde{X}_t^{rs,r's} = \left(\frac{\Gamma^{rs}}{\Gamma^{r's}} \frac{P_t^{r's}}{P_t^{rs}}\right)^{-\theta^s} \tag{57}$$

where we have substituted for $\phi_t^{rs,r's} = \left(\frac{A_t^{r's}}{\tau_t^{rs,r's} x_t^{r's}}\right)^{\theta_s}$.

Using data on transport costs from above, we can construct the left-hand side, introduce a mean-zero stochastic term $\epsilon_t^{rs,r's}$, and postulate a log-linear stochastic gravity-type regression with fixed effects only of the form:

$$\ln \tilde{X}_t^{rs,r's} = \alpha_t^{rs} + \beta_t^{r's} + \epsilon_t^{rs,r's}. \tag{58}$$

Using data on θ^s, we can recover estimates of $\left(\frac{P_t^{rs}}{\Gamma^{rs}}\right) = exp\left(\widehat{\alpha_t^{rs}}\right)^{-1/\theta^s}$. Note that this al-

[35] See equation (12) in Eaton and Kortum (2002).

lows us to obtain relative intermediate price indices across regions within a sector. This is, however, insufficient for our purposes as we have to make these price indices comparable across sectors. To establish a common anchor for the sector-level price indices across regions, we use data on relative intermediate input price differences across sectors from the GGDC Productivity Level Database and scale all derived estimates in a given sector using these data. Finally, we use the sector-region-specific price indices[36] to aggregate them up to the sector-country-specific price of the intermediate input bundle which allows us to calculate the intermediate input quantity, M_t^{rs}, from expenditure on intermediate inputs using $\prod_{s'=1}^{S}\left(\frac{P_t^{rs'}}{\gamma^{rs',rs}}\right)^{\gamma^{rs',rs}}$. With all data at hand, we calculate the region-sector-specific $\ln(A_t^{rs})$:

$$\ln(A_t^{rs}) = \ln(Y_t^{rs}) - \ln\left(\frac{P_t^{rs}}{\Gamma^{rs}}\right) - \gamma^{rs}\xi^{rs}\ln(H_t^{rs}) - \gamma^{rs}\left(1-\xi^{rs}\right)\ln(L_t^{rs}) - \left(1-\gamma^{rs}\right)\ln(M_t^{rs}) + 1/\theta^{s}\ln(\pi_t^{rs,rs}), \tag{59}$$

where γ^{rs} is the labor-cost share, θ^s the trade elasticity, $\pi_t^{rs,rs}$ is the local absorption share obtained from the constructed trade flow matrix, Y_t^{rs} are aggregated gross revenues in r,s, L_t^{rs} are total hours worked in r,s, and M^{rs} is obtained from total input expenditures in r,s divided by the intermediate input price as defined above.

Panel (a) of Figure 6.2 reports the distribution of measured TFP, A_t^{rs}, across French regions aggregated over sectors using employment-based weights. According to the figure, the south of the country (where manufacturing activities are predominant) and Île de France, the region of Paris (the center of financial services), emerge as the areas with the highest productivity levels.

E.7 Estimation of the stochastic process for TFP

While our model is silent on the origins of TFP, we model its level as a function of time-invariant components that vary at the sector and region level, yearly variations that are common across France, and a component that is a function of labor employed in a region sector. The latter takes into account considerations that TFP might not be purely deterministic, but is to a certain extent an outcome of the pre-determined agglomeration of labor L_t^{rs} in a region-sector.[37] For our analysis of the role of uncertainty, we are particularly interested in the region-sector-time stochastic component, η_t^{rs} of measured TFP, that is not captured by the above elements of the TFP process. To isolate the stochastic part η_t^{rs} but at the same time keep the level of TFP varying across sector, we exclude any variation in TFP that is common across France and time varying (captured by time fixed-effects) and

[36] Note that we calculate $\Gamma^{rs} = \Gamma\left(1+\left(1-\eta^{rs}/\theta^{s}\right)\right)$ from elasticities to isolate the level of the price index.

[37] Note that this is in line with the model where we consider a deterministic evolution of the distribution of the labor force, such that labor allocated to a certain region and sector at time t is the deterministic consequence of the distribution of labor and TFP realizations at time $t-1$. Therefore, through the lens of the model, the labor allocation L_t^{rs} is exogenous to the actual realization of technological shocks at time t.

the pre-determined agglomeration of labor:[38]

$$\ln A_t^{rs} = \kappa \ln L_t^{rs} + \delta_r + \delta_s + \delta_t + \eta_t^{rs}, \tag{60}$$

where κ is the elasticity of measured TFP to labor allocation, δ_r, δ_s, and δ_t are, respectively, region-, sector-, and year-fixed effects. The fraction of total variation in measured TFP explained by these deterministic components is 78.6%. The elasticity of labor allocation is small but statistically significant $\hat{\kappa} = 0.028$, which explains somewhat less than 1% of the total variation in log TFP. We define the stochastic component of TFP as the region- and sector-specific components of measured TFP plus the residual in (26):

$$\ln a_t^{rs} = \delta_r + \delta_s + \eta_t^{rs}. \tag{61}$$

Panel (b) in Figure 6.2 reports the distribution of the stochastic component of TFP, a_t^{rs}, across French regions aggregating using sector-level employment shares as weights. Differences with respect to the overall measured TFP in panel (a) of the figure are modest in magnitude. Generally, the pattern that emerges is a shorter range of variation of the stochastic component. Thus, considering the stochastic component of TFP is a more conservative approach for our purposes. In fact, it reduces the spectrum of realizations of the aggregate stochastic shock that agents are uncertain about, thus, shrinking the room for uncertainty to play a role.

Using the panel data on $\ln a_t^{rs}$ for the years 2003-2014, we postulate and estimate an autoregressive process of the form that informs the uncertainty around realized TFP levels that we are going to feed into the quantification exercise:

$$\ln a_t^{rs} = \mu^{rs} + \rho \ln a_{t-1}^{rs} + \iota^{rs} \varepsilon_t^{rs}, \tag{62}$$

where $\varepsilon_t^{rs} \sim i.i.d.(0,1)$, ρ is a common autoregressive coefficient, the region and sector specific parameter μ^{rs} scales the mean of the process, and ι^{rs} parametrizes the region-sector specific volatility that tunes the degree of uncertainty in region r and sector s. As long as $|\hat{\rho}| < 1$, the process is stationary. Then, $\ln a^{rs} \equiv \hat{\mu}^{rs}/(1-\hat{\rho})$ and $\sigma^{rs} \equiv \hat{\iota}^{rs}/\sqrt{1-\hat{\rho}^2}$ are, respectively, the asymptotic long-run mean and standard deviation of the stochastic component of TFP in region r sector s.

Table 12 reports the moments of the process for $\ln(a_t^{rs})$, where the estimate of the autoregressive coefficient is $\hat{\rho} = 0.503$, and the volatility $\hat{\iota}^{rs}$ is estimated as the squared root of the average of squared residual based on (28). Panel (a) of Figure 6.3 reports the distribution of the weighted long-run mean, $\ln a^{rs}$, across French regions, and panel (b) reports the

[38] The model assumes a Markovian stationary stochastic process for the aggregate stochastic shock. Thus, only the last event in the history of shocks matters, not the entire history. This implies that (pre-determined) labor allocations before time t should not play a role. Furthermore, among production factors other than labor, note that the purchase of materials at time t is a general equilibrium outcome, endogenous to the contemporaneous realization of technological shocks; instead, the stock of structures is constant over time. Consequently, taking into account the contemporaneous allocation of labor (other than region, sector and time fixed effects) is the most conservative specification that is also admissible through the lens of the model.

Table 12: Summary statistics for the dynamic process of $\ln(a_t^{rs})$

	mean	std.dev.	p10	p50	p90
Long run mean $\ln a^{rs}$	2.436	0.108	2.318	2.438	2.543
Long run standard deviation σ^{rs}	0.056	0.058	0.015	0.037	0.117

weighted long-run standard deviation, σ^{rs}, where weights reflect sector-level employment in each region.

E.8 Intertemporal equilibrium parameters

For the intertemporal equilibrium, three parameters are of particular importance: ν (the dispersion of idiosyncratic tastes for jobs); β (the discounting parameter); and ζ^{jn} (the mobility costs of switching from job j to n.

To estimate these parameters, we rely on the French administrative employer-employee dataset *Declaration Annuelle des Donnees Sociales* published by INSEE. This dataset provides information on the employees and their wage in a given job n (i.e., region r, sector s, and occupation/skill level k) in a given year t. Moreover, it contains information on the same employees on their job j in the previous year $t-1$ as well as the wage earned then. Using data for the pair of consecutive years of 2009-2010 and 2012-2013, we can compute transitions between every pair of jobs j and n as well as associated wages for about mm. 27 workers per year.

E.9 Estimation of moving costs ζ^{jn}

We estimate moving costs based on the methodology proposed by Artuç and McLaren (2015) that is based on equation (3). We calculate the implied moving costs directly from data on movers over the period 2009-2010 as well as 2012-2013. Specifically, we consider gross migration flows between the consecutive pair of mentioned years, t and $t+1$, and pair of jobs of origin j and destination i, $m_t^{aj,i}$ for three types of age groups (21-25 years, 26-45 years, and 46-65 years). Equation (3) multiplied by the number of workers in each job, $l_t^{a,i}$, motivates the following estimating equation. First, write the right-hand side of the equation for movers as an exponential of a log-linear index. In that index, consider $\psi_t^{a,i} \equiv \frac{1}{\nu}\beta\mathbb{E}_t\left[\bar{V}^{ai}\left(L_{t+1}, Z_{t+1}\right)\right]$ and $\phi_t^{a,j} \equiv \ln\sum_{k\in J} e^{\frac{1}{\nu}\left(\beta\mathbb{E}_t\left[\bar{V}^{ak}(L_{t+1},Z_{t+1})\right]-\zeta_t^{aj,k}\right)}$ as fixed age-job-pair fixed effects specific to t. Then, after introducing a mean-zero stochastic term $\varepsilon_t^{aj,i}$, we can re-express Equation (3) as an exponential-form estimation equation that can be estimated using a Poisson Pseudo-maximum Likelihood (PPML) estimator:

$$m_t^{aj,i} l_t^{a,i} = movers_t^{aj,i} \quad = \quad exp(\psi_t^{a,i} + \phi_t^{a,j} - \zeta_t^{aj,i}/\nu + \varepsilon_t^{aj,i}). \tag{63}$$

We obtain an estimate of $\zeta_t^{aj,i}/\nu$ as the residual of a fixed-effects PPML regression. Since this residual is only identified for all job-pairs with non-zero flows, we explain these moving

Table 13: Linear regression explaining job-to-job moving costs

	$\widehat{\zeta^{aj,i}_{2009}/\nu}$	$\widehat{\zeta^{aj,i}_{2012}/\nu}$
log(Distance)	1.270***	1.260***
	(0.0102)	(0.0109)
Switch Skill	4.362***	4.779***
	(0.0240)	(0.0240)
Constant	5.970***	4.288***
	(0.0114)	(0.0109)
Fixed effects	Sector-pair	Sector-pair
Observations	9,442	8,864

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

costs using the following regression

$$\widehat{\zeta^{aj,i}_t/\nu} = \beta_1 \ln(Distance^{rr'}) + \beta_2 \mathbf{I}(k \neq k') + \beta_{ss'} + \varepsilon^{aj,i}_t,$$

and predict moving costs for those cells with unobserved movers using the log great-circle distance between the centroids of regions r and r' (being zero for $r = r'$), $\ln(Distance^{rr'})$, an binary indicator for switching skills and sector-pair fixed effects. The latter step of explaining moving costs with observables allows us to obtain distinct moving costs for a larger set of the pairs of $2,156$ jobs from one year to another. We present the regression for the moving costs in Table 13.

In order to obtain an estimate of ν, we follow Artuç and McLaren (2015) and use the structural interpretation of the fixed effects to isolate $1/\nu$:

$$\phi^{a,j}_t = \beta/\nu \mathbb{E}_t \left[\bar{V}^{aj}(L_{t+1}, Z_{t+1})\right]$$

and use the estimates of the first stage along with information on job-year-specific real wages to estimate the following regression:

$$\Psi^{a,i}_t = \alpha^a_0 + 1/\nu \times \beta \times u(w^i_t/P^i_t) + \varepsilon^{a,i}_t,$$

where $\Psi^{a,i}_t = \phi^{a,i}_t + \beta \times \psi^{a,i}_t - \beta \times log(L^{i,a}_{t+1})$ and real wages are instrumented using lagged values. Moreover, we fix the discounting parameter at a value of $\beta = 0.95$. Real wages are obtained from the average job-age spell-specific nominal wage rates reported in the employer-employee data and the price indices of output estimated before and aggregated using sectoral consumption shares as weights.

Table 3 provides descriptive statistics of the estimated moving costs. Job-to-job flows

between 2012 and 2013 indicate that, on average, 98.2% of the agents in a given job do not move to other jobs from one year to another. Overall, moving costs are smaller for changing occupation only within the same region and sector. In the subsample below the 10th percentile of the moving-cost distribution, moving costs are higher for changing the sector and they are slightly higher for changing the region. In the rest of the distribution, mobility frictions increase when changing the sector, while the cost of moving between regions or occupations remain flat. To interpret the magnitude of moving costs, rearrange the policy (3)

$$moving\ cost^{jn} = \beta \mathbb{E}[value^{n} - value^{j}] - \nu \ln \left(\frac{movers^{jn}}{stayers^{jj}} \right),$$

take a steady-state perspective, in which certainty equivalence holds and the value of a job is given by the current utility divided by the discount rate $1 - \beta$

$$moving\ cost^{jn} = \frac{\beta}{1-\beta}(utility^{n} - utility^{j}) - \nu \ln \left(\frac{movers^{jn}}{stayers^{jj}} \right),$$

and finally make use of the log-utility assumption, to substitute for real wages:

$$moving\ cost^{jn} = \frac{\beta}{1-\beta} \ln \left(\frac{real\ wage^{n}}{real\ wage^{j}} \right) - \nu \ln \left(\frac{movers^{jn}}{stayers^{jj}} \right).$$

Given $\beta = 0.95$, $\nu = 4.5281$ and a fraction of stayers equal to 98.2% on average, moving costs are as large as several times the log difference in observed real wages between origin and destination in the data, and this level is further inflated by subtracting the log of a small ratio of movers to stayers.

Two remarks are worthy. First, PPML estimation identifies the dispersion of moving costs from observed flows. Although the method naturally accommodates zeros, the estimated levels of moving costs are informative the greater the observed fraction of movers that correspond to the lower percentiles of the distribution. Second, in the estimation (and in our description above) we did not consider differences in moving costs by age spell. In this way, we load all the differential behavior between age cohorts to endogenous channels of the model only (old agents have a lower continuation value due to a shorter expected lifetime time) and not to exogenous differences in moving costs.