Technology Diffusion, Worker Mobility and the Returns to Skill☆

January 2015
First Version: January 2013

Damir Stijepic
Goethe University Frankfurt

Abstract

In this paper I illustrate how the diffusion across firms of a skill-neutral technology leads to a skill-biased impact on the economy. The model identifies *(i)* differences in inter-firm mobility between skill groups, *(ii)* productivity dispersion across firms within industries, and *(iii)* differences in wages between small and large firms as key determinants of the skill premium. Calibrated to match differences in inter-firm mobility between skill groups and rising productivity dispersion across firms, the model ascribes one-third of the sharp increase in the skill premium in U.S. manufacturing from 1977 to 1997 to skill-neutral technical progress and the technology diffusion process itself. Technical progress complementing high-skill workers accounts for three-fifths of the increase in the skill premium.

Keywords: wage inequality, technical change, heterogeneous firms, frictional labor markets, on-the-job search, sorting, versatility
JEL: J31, J62, I24, I26, O33

☆I thank Jennifer Abel-Koch, Alexander Bick, Emmanuele Bobbio, Ester Faia, Guido Friebel, Nicola Fuchs-Schündeln, Peter Funk, Denis Gorea, Nils Gornemann, Philipp Harms, Boyan Jovanovic, Philipp Kircher, Sebastian Krautheim, Joachim Möller, Friedrich Poeschel, Sigrid Röhrs, Thorsten Schank, Jakob Schwab, Ctirad Slavík, Denis Stijepic, Fraser Summerfield, Hitoshi Tsujiyama, Klaus Wälde, the participants of the Midwest Macro Meeting (Urbana-Champaign, Spring 2013), the European Economic Association Meeting (Göteborg, 2013), the meeting of the Verein für Socialpolitik (Düsseldorf, 2013), the Midwest Economic Theory and Trade Conference (Indianapolis, 2014), and the Cologne Workshop on Macroecnomics (Köln, 2014) for helpful comments. Furthermore, I thank the Bureau of the Census, especially Asheleigh Perry, for assistance with the data. I gratefully acknowledge financial support from the Excellence Cluster "Normative Orders" at the Goethe University, the "Forschungstopf" of the Department of Economics and Business Administration at the Goethe University, and the Goethe Money Macro Association. The usual disclaimer applies.

Email address: `mail@damir.stijepic.com` (Damir Stijepic)

1. Introduction

Skill-biased technical change is perhaps the most prominent explanation for the rise in wage inequality in the United States over the last decades. While there is an extensive literature studying the implications of the advancement of skill-biased technologies, little attention has been devoted to the effect of the technology diffusion process itself on wage inequality. In particular, the link between heterogeneity across firms with respect to the adoption of a new technology and the skill premium remains mostly unexplored. I fill this gap by illustrating in a new framework how the adoption of a skill-neutral technology leads to a skill-biased impact on the economy. While the model is consistent with well-known stylized facts, it generates distinctive predictions that are again in line with the data. In particular, the model postulates a close link between the skill premium and the differential firm size wage premium between skill groups. Quantitatively, the model ascribes one-third of the sharp increase in the skill premium in U.S. manufacturing from 1977 to 1997 to skill-neutral technical progress and the technology diffusion process itself.

I consider a competitive industry model à la Hopenhayn (1992). Firms endogenously select in an industry. The equilibrium distribution of firm productivity, employment, and output is endogenously determined by firms' profit maximizing decisions.[1] I depart from the assumption of perfect competition in all factor markets by introducing frictional labor markets à la Cahuc et al. (2006). Informational frictions hinder the allocation of workers to the most productive firms, and wages are bargained for. I then extend the framework in two important aspects. First, I allow for technical change in the sense that there is an ex post technology choice in addition to ex ante firm heterogeneity. Second, workers differ in the tasks they are able to perform and, similarly, jobs differ in their task requirements. Therefore, it is not sufficient for firms and workers to overcome informational frictions in order to produce. A mismatch between the task required for the job and the tasks the worker is able to perform may still prevent the formation of a worker–firm match. Worker flows across firms are determined by both frictional and structural factors in the resulting framework.

While I study the impact of technical change on the skill premium, I focus on one specific dimension of skill, i.e., versatility. Workers may increase through education the array of tasks they are able to perform. However, all tasks are

[1] Hopenhayn's (1992) framework also features firm productivity dynamics, which are not modeled here.

equally valuable in production. The only advantage from versatility is the flexibility resulting from the ability to switch between more jobs. The key mechanism exploited in this paper relies on differences in versatility across worker groups. A higher versatility allows workers to be more mobile between firms since, for instance, an efficient reallocation is less likely to be hindered by unmet job requirements. It is this mobility advantage that already gives rise to a positive skill premium in the model.[2]

This paper's main contribution is a new link between technical change and the skill premium. Insofar as technical change increases the dispersion of productivity among firms, it exerts an upward pressure on the skill premium. High-skill (or high-versatility) workers' inter-firm mobility advantage is more pronounced in a high-dispersion environment. Therefore, high-skill workers' relative wages are likely to rise. Intuitively, if firms are similar in productivity, the returns from switching firms are low. However, if the disparities between firms are substantial, so will be the returns. Wage differences across worker groups who differ in inter-firm mobility are amplified.

Empirically, I provide evidence in favor of the model's microstructure, show that key patterns of the proposed link between productivity dispersion and the skill premium are observed in the data, and illustrate the quantitative importance of the channel in a numerical exercise. Specifically, I show that statistics on employer–employer transitions, transitions into unemployment, and occupational changes obtained from the Current Population Survey Basic Monthly data are in line with the modeling assumptions. Furthermore, I provide evidence that a key prediction of the model, i.e., a close relation between the skill premium and the differential firm size wage premium between skill groups, is indeed in line with the data. Finally, I show that the calibrated model ascribes one-third of the sharp increase in the skill premium in U.S. manufacturing from 1977 to 1997 to skill-neutral technical progress and the technology diffusion process itself.

This paper is related to the literature on skill-biased technical change. See Acemoglu and Autor (2011) for a recent assessment and further reading. The canonical model assumes two distinct skill groups that perform two different and imperfectly substitutable tasks or produce two imperfectly substitutable goods. Technology is assumed to take a factor-augmenting form, which, by complement-

[2]Lise and Postel-Vinay (2014) develop a multi-dimensional sorting model: workers differ in skills along several dimensions, jobs require mixes of various types of skills, and workers improve skills that they regularly use.

ing either high- or low-skill workers, can generate skill-biased demand shifts. Acemoglu (2002) acknowledges the endogeneity of technical progress and analyzes the factors determining its direction and bias. Caselli (1999) focuses on substitutability among technologies. Machine-specific skills are needed to operate machines. The acquisition of such skills is costly and workers are heterogeneous in the cost of acquisition. A technology revolution is skill-biased if the new skills are more costly to acquire than the skills required by preexisting equipment. Aghion et al. (2002) stress the general purpose nature of the new information technologies in contrast to occupation- or industry-specific technologies. Workers accumulate skills through learning-by-doing. A more general technology allows a larger degree of transferability of skills across the different sectors of the economy. Therefore, adaptable workers, i.e., workers who are productive with the new technology, preserve more skills when moving to the leading-edge sector and the wage premium of adaptable workers rises.

My contribution differs conceptually from the aforementioned literature. The mechanism I propose does not rely on any complementarity between technology and skill. Specifically, at a given firm, low-skill and high-skill workers may be in general equally efficient at operating any of the available technologies. Nor do I assume that high-skill workers are able to adapt to new technologies faster or better. It is differences in inter-firm mobility, which are in turn motivated by differences in versatility, that are driving the skill-biasedness of technical progress in this model. As will be evident from the subsequent analysis, inter-firm mobility, by affecting the degree of competition between firms, is related to two core economic issues: the allocation of resources across economic activities and the distribution of income across factors of production.[3]

My approach is well in line with empirical studies that highlight the role of establishment- or firm-specific wage premia in generating the recent increases in wage inequality. For instance, Card et al. (2013) fit linear models with additive person and establishment fixed effects à la Abowd et al. (1999) for West Germany for the years 1985–2009. They estimate that the rise in the variance of the person component of pay contributes about 40 percent of the overall rise in the variance of wages, the rise in the establishment component contributes around 25 percent, and

[3]Stijepic (2015a) develops a heterogeneous firm model of intra-industry trade with limited inter-firm mobility of workers in order to study the impact of international trade on wage inequality. Trade openness *(i)* amplifies disparities in profitability between small and large firms, *(ii)* raises within-group wage inequality, and *(iii)* increases wage differentials between worker groups who differ in inter-firm mobility.

their rising covariance contributes about a third. Furthermore, they find that two-thirds of the increase in the pay gap between higher- and lower-educated workers are attributable to a widening in the average establishment pay premia received by different education groups. Increasing workplace heterogeneity and rising assortativeness between high-wage workers and high-wage establishments likewise explain over 60 percent of the growth in inequality across occupations and industries. See Andersson et al. (2012) for a study of the U.S. labor market.[4]

This paper also complements existing models of matching and sorting in the labor market. Specifically, it is related to the literature that stresses the importance for workers of occupational matching (e.g, Kambourov and Manovskii, 2009; Kircher et al., forthcoming), firm matching (e.g, Jovanovic, 1979; Burdett and Mortensen, 1998; Alvarez and Shimer, 2011), or both occupational and firm matching (e.g, Papageorgiou, 2010; Kramarz et al., 2014). However, this branch of the literature does not address how versatility, in the sense of being able to perform a wider range of tasks, and productivity dispersion across firms affect the skill premium.

The paper is structured as follows. In Section 2 I present stylized facts to motivate the model. The formal exposition of the model is in Section 3. I characterize the equilibrium of the model in Section 4. The quantitative exercise is in Section 5. Section 6 draws some conclusions.

2. Stylized Facts

In this section, I analyze the differences in mobility patterns between skill groups, i.e., the frequency of occupational changes, employer–employer transitions, and separations into unemployment. Furthermore, I describe the evolution of productivity dispersion and wage inequality in the United States over the last decades with a particular emphasis on the differences between establishment size classes.[5] I focus here on manufacturing since this industry is traditionally well covered. Production and non-production workers serve as proxies for low-skill

[4] Various empirical studies have documented that workers using new technologies are substantially better paid than nonusers. However, the new technology workers are typically already better paid before entering the new technology jobs. For instance, Entorf et al. (1999), relying on French linked employer–employee data for the early 1990s, find that *(i)* computer users enjoy a wage premium of 15 to 20 percent relative to nonusers, but that *(ii)* an individual worker's wage increases by less than two percent over the course of two to three years when entering a new technology job.

[5] In the model I use firms as the unit of analysis and consider statistics according to firm size classes. The data, on the other hand, is available for establishment size classes.

and high-skill workers, respectively, since this information is consistently available over the entire sample period. A description of the data sets is in Appendix A. While the stylized facts presented in this section are of interest in their own right, the primary purpose is to motivate the paper's theoretical contribution. In this section I also sketch key aspects of the model in light of the stylized facts. However, the formal exposition of the model is in Section 3.

2.1. Mobility Patterns

In this section I document monthly changes in the employment status of production and non-production workers in U.S. manufacturing for 1996–2009 based on the Current Population Basic Monthly data. Specifically, I consider three categories of changes in employment status: employer–employer transitions, separations into unemployment, and activity or duty changes of workers who stay with the same employer.[6] Following Fallick and Fleischman (2004), I exploit the dependent interviewing techniques, employed by the Bureau of the Census since January 1994, to identify employer–employer transitions. I rely on self-reported activity or duty changes. However, using the U.S. Census occupational classification system instead still yields similar results.[7] See Appendix A for further details.

Figure 1 decomposes changes in employment status into the aforementioned three categories for both production and non-production workers in U.S. manufacturing. Employer–employer transitions account for 22 and 25 percent of overall changes among production and non-production workers, respectively. While the difference in the share of employer–employer transitions is relatively small between skill groups, there is considerable heterogeneity in the shares of activity changes and separations into unemployment. For production workers, the share of activity changes in overall changes is 36 percent and the share of separations into unemployment 42 percent. For non-production workers, the respective shares are 47 percent and 29 percent. In summary, conditional on a change in employment status, non-production workers are more likely to switch employers, more

[6] The employer–employer transitions category encompasses also separations into self-employment, and the separations into unemployment category includes separations into inactivity. Therefore, the subsequent decompositions are comprehensive in the sense that they capture all reported employer–employee separations.

[7] The mobility patterns presented in this section are robust in various respects. They hold for skill groups based on educational attainment, for job-changes into less paid occupations, and in booms and busts. Statistics are available upon request.

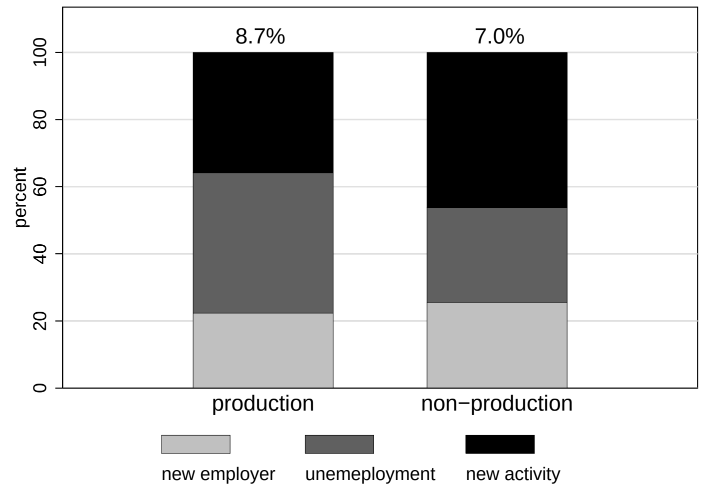

Figure 1: Decomposition of monthly changes in employment status of production and non-production workers employed in U.S. manufacturing for the years 1996–2009. Top bar values denote the share of workers experiencing a change in employment status in percent of total employment. Author's calculations based on the Current Population Survey Basic Monthly data as provided by the National Bureau of Economic Research (`http://www.nber.org/data/cps_basic.html`). See Appendix A for further details.

likely to change their activity or duties while staying with the same employer, and less likely to separate into unemployment than are production workers.

The differences in mobility between skill groups are at the core of the model presented in the next section. I attribute these differences to differences in versatility. In a nutshell, high-skill workers are able to perform a wider range of tasks, jobs differ in task requirements, and task requirements may change while employed at a firm. If the task requirement changes, high-skill workers are more likely to be able to perform the new task, whereas low-skill workers, being less versatile, are more likely to separate into unemployment as a result of a task mismatch, i.e., a mismatch between the task required for the job and the tasks the worker is able to perform. Therefore, high-skill workers are less likely to separate into unemployment and more likely to change tasks while staying with the same employer. Furthermore, worker–employer matches require in this setup firms and workers not only to overcome informational frictions but also an overlap in the job requirements and the tasks the worker is able to perform. Therefore, fewer employer–employer transitions are prevented by unmet job requirements among high-skill workers since high-skill workers are able to perform a wider array of tasks.

The relevant measure of inter-firm mobility in on-the-job search models is not the extent of employer–employer transitions alone, but rather employer–employer transitions relative to separations into unemployment. Intuitively, separations into unemployment represent negative mobility shocks. The more pronounced the separations shocks, the less likely are individuals to allocate to a specific job. Therefore, employer–employer transitions are to be scaled by separations into unemployment. See Stijepic (2015b) for further details. Furthermore, the ratio of employer–employer transitions relative to separations into unemployment is related to key concepts of the model presented in the next section. Finally, note that the ratio is 0.53 among production workers and 0.86 among non-production workers. Therefore, non-production workers indeed exhibit a higher degree of inter-firm mobility than production workers according to the given measure.

Stijepic (2015b), relying on the Survey of Income and Program Participation, studies the determinants of inter-firm mobility. In particular, he also accounts for an individual's versatility using a direct measure based on the number of different courses attended in high school. Stijepic (2015b) finds *(i)* a strong positive correlation between a worker's education and versatility, and *(ii)* a substantially higher inter-firm mobility among versatile workers even after controlling for an extensive set of covariates. Specifically, individuals with above-median versatility are 1.43 times likelier to switch employers than to separate into unemployment relative to individuals with below-median versatility. The effect of versatility on inter-firm mobility is, therefore, of a similar magnitude as the effect of a college degree on a high school dropout's inter-firm mobility.

2.2. Productivity and Wage Dispersion

Dunne et al. (2004) exploit establishment level data to investigate the relation between the dispersion of wages and the dispersion of labor productivity across establishments in U.S. manufacturing. They find that the between-plant wage and productivity dispersion increased substantially from 1975 to 1992, and that "virtually the entire increase in overall dispersion in hourly wages for U.S. manufacturing workers from 1975 to 1992 is accounted for by the between-plant components" (Dunne et al., 2004, pg. 399). Furthermore, the authors stress that these trends occur mostly within industries and are not a between-industry phenomenon. And finally, they find that a significant fraction of the rising dispersion in wages and productivity is accounted for by changes in computer investment across plants.[8]

Less known is that labor productivity dispersion across establishment size classes increased as well. Figure 2 depicts the value added per worker in four establishment size classes scaled by the overall value added per worker in the respective year for U.S. manufacturing. The figure shows a substantial increase in labor productivity disparities between large and small establishments in particular since the 1970s. For instance, the value added per worker at establishments with at least one thousand employees was 1.4 times higher than at establishments with at most one hundred employees in 1954. The ratio was still 1.4 in 1972. Thereafter, the gap in value added markedly opened up and the ratio amounted to 2.3 in 1997.

The literature highlights the information and communication technology revolution that started essentially in the 1970s as a potential source for the rise in disparities in value added per worker across establishment and firm size classes. Increases in organizational size are usually associated with increasing complexity and increasing problems in communication and coordination. Therefore, advances in information and communication systems are likely to be predominantly beneficial to large organizations. Furthermore, the adoption of administrative innovations, of which information systems are an example, are positively related to organizational size insofar as economies of scale can be realized. For instance, larger organizations can typically spread the fixed costs of implementing an information system over a larger base.[9] Lee and Xia (2006) report in their meta-study a mean correlation of 0.2265 between organizational size and the adoption of information technology with the 95 percent confidence interval ranging from 0.2073 to

[8] It is often argued that for technical change to be a compelling explanation for the rise in wage inequality, trends have to be similar across different countries having access to the same technology. The work by Faggio et al. (2007, 2010) points in this direction. Using panel data on UK firms over the 1984 to 2001 period, Faggio et al. (2007, 2010) obtain following results: First, the vast majority of the increase in individual wage inequality is a between-firm phenomenon and most of the growth of wage and productivity dispersion is within industries. Second, the increase in firm-level productivity dispersion is mainly in the service sector of the economy suggesting that studies based on manufacturing alone underestimate the rise of economy-wide productivity dispersion. Third, the increase in labor productivity dispersion is mainly driven by an increase in total factor productivity dispersion rather than being solely due to an increase in the dispersion of capital-labor ratios. Fourth, those industries that had the most rapid increase in the use of information and communication technology also had the most rapid increase in productivity dispersion. And finally, they find similar patterns for Norway and France, however, less pronounced.

[9] See, e.g., Gremillion (1984) for a more detailed exposition.

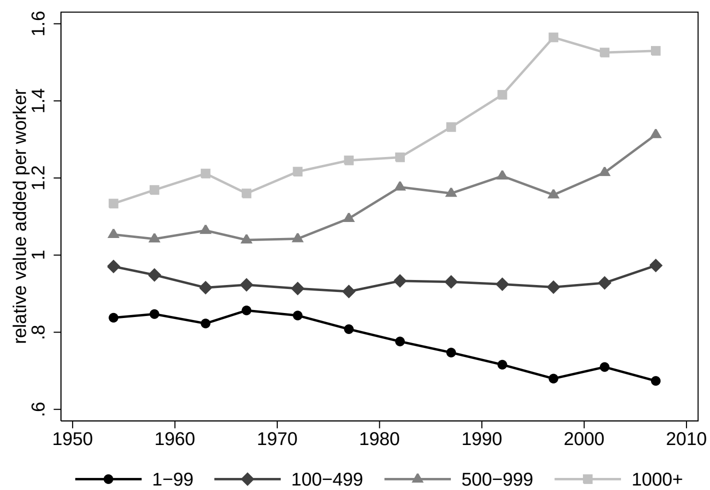

Figure 2: Value added per worker in U.S. manufacturing by establishment size class scaled by the overall average in the respective year. Establishment size defined in terms of the average number of workers. Author's calculations based on Census of Manufactures reports. See Appendix A for further details.

0.2457.[10] While I focus on the information and communication technology revolution in this paper, other factors are likely to have contributed to the widening of value added per worker differentials across establishment size classes as well. For instance, it is a well-known stylized fact that larger firms a more likely to export and, therefore, to profit from access to foreign markets. Furthermore, trade integration by expanding the size of the market encourages firms to innovate and to adopt new technologies. Hence, globalization is likely to increase productivity at large firms relative to small firms (e.g., Tybout, 2008; Melitz and Trefler, 2012).

The skill premium has been increasing in the United States over the last decades (see, e.g., Acemoglu, 2003). Less known is that the increase in the skill premium was accompanied by an increase in high-skill workers' establishment size wage premium relative to that of low-skill workers. Figure 3 depicts the skill premium and the differential size premium between non-production and production workers in U.S. manufacturing, where the size premium is defined as the wage premium enjoyed by workers at establishments with at least 500 employees relative

[10]Lee and Xia (2006) document substantial heterogeneity across studies, where the obtained correlations range from -0.300 to 0.570. They identify in particular type of innovation, type of organization, stage of innovation adoption, and scope of adoption as important factors affecting the size–adoption relation.

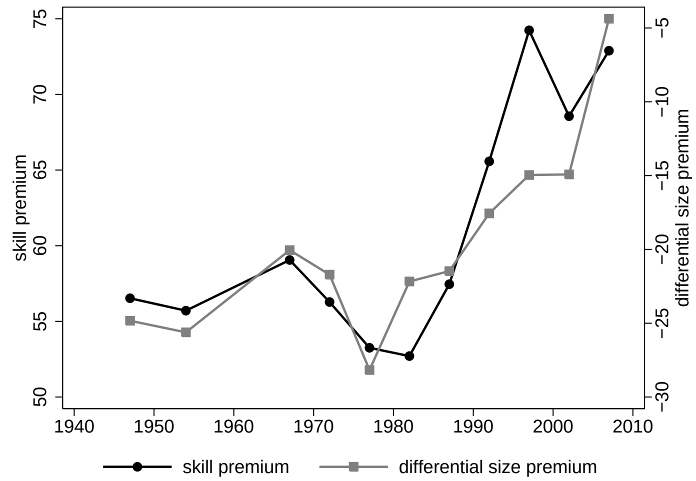

Figure 3: Skill premium and differential size premium between skill groups in U.S. manufacturing. The size premium is defined as the wage premium enjoyed by workers at establishments with at least 500 employees relative to workers at establishments with less than 500 employees. Non-production workers and production workers serve as proxies for high- and low-skill workers, respectively. Author's calculations based on Census of Manufactures reports. See Appendix A for further details.

to workers at establishments with less than 500 employees.[11] The skill premium increased by 17 percentage points from 56 percent to 73 percent over the sample period, where most of the increase occurred in the 1980s and 1990s. The difference in the establishment size wage premia increased from -24 percent to -4 percent over the same time span. This suggests that the skill premium is related to differences in wages between small and large establishments.

Intuitively, the comovement of the skill premium and the differential firm size wage premium between skill groups is generated in the model as follows. It is in particular the large firms, which are on average more productive and pay higher wages, that profit from technical progress. Low- and high-skill workers differ in versatility, i.e., the range of tasks they are able to perform. High-skill workers' higher versatility translates into a higher degree of inter-firm mobility. A higher degree of inter-firm mobility in turn intensifies the competition between firms and allows high-skill workers to appropriate a larger share of the surplus that is generated at large firms through the adoption of a more advanced technology. There-

[11] Note that establishments with less than 500 employees account on average for around 60 percent of overall employment in manufacturing over the sample period.

fore, both the firm size wage premium and wages of high-skill workers increase relative to those of low-skill workers.

Beyond this rent share effect, there is an allocation effect. The differences in inter-firm mobility between skill groups are also reflected in differences in the distribution of workers over firm size classes. Intuitively, high-skill workers, being more mobile, are more likely to find and to match with the most productive and at the same time large firms. Therefore, they represent a disproportionately large share of the workforce at the respective firms. Since the gains from technical progress are mostly realized at large firms, high-skill workers also profit disproportionately from technical change.

3. Framework

The framework is closely related to Hopenhayn's (1992) model of endogenous selection of heterogeneous firms in an industry and Cahuc et al.'s (2006) on-the-job search model with outside-offer matching and bargaining. However, I extend the models along several dimensions. In addition to an exogenous ex ante distribution of entrepreneurs' managerial skills, which affects the productivity of firms, there is an ex post technology choice. Therefore, there is an additional endogenous component to the distribution of productivity. Furthermore, entrepreneurs decide on how many production facilities to set up, which endogenizes the sampling distribution. Finally, I introduce additional heterogeneity across production facilities and workers in the sense that production facilities may differ in the tasks required for production, and, similarly, workers may differ in the tasks that they are able to perform. Therefore, worker flows are not determined by informational frictions alone but are affected by the heterogeneity in tasks across production facilities and workers as well. One implication is that the model exhibits both frictional and structural unemployment, in contrast to the canonical on-the-job search model.

3.1. The Economy

I consider an economy that consists of a single sector with one homogeneous multipurpose good, which serves as the numeraire. There is a mass M of risk-neutral workers. Furthermore, there is a continuum of tasks, t, in the economy, that I normalize to unity. All workers are able to perform a share, $\alpha_L \in (0,1)$, of the tasks, where workers are uniformly distributed over tasks. By incurring an education cost of $f_g \overset{i.i.d}{\sim} \Psi$, workers can increase the share of tasks they are able to perform to $\alpha_H \in (\alpha_L, 1)$ before entering the market.[12] Workers who are able to perform only a share α_L of the tasks are referred to as low-skill workers; workers who are able to perform a share α_H of the tasks are referred to as high-skill workers. The mass of low-skill workers is denoted by M_L, and that of high-skill, by M_H. Once a worker enters the market, all eventually incurred costs are assumed to be sunk.

There is a pool of potential entrepreneurs, who can choose to enter the market by incurring a cost of $f_e > 0$. Let $\tilde{N}$ denote the equilibrium mass of entrepreneurs who decide to enter the market. On incurring the entry cost, entrepreneurs observe their managerial skills, $s \overset{i.i.d}{\sim} \tilde{\Gamma}$ for $s \in [\underline{s}, \overline{s}]$, and the entrepreneur-specific cost, $f_a \overset{i.i.d}{\sim} \Phi$, of adopting the new and more advanced technology $a_n > 1$.[13] I normalize the old and less advanced technology to unity, i.e., $a_o = 1$. Before production starts, the entrepreneur decides whether or not to participate in the market, whether or not to adopt the more advanced technology, and how many production facilities, n, to set up.[14] Entrepreneurs may participate in the market at a cost of $f_p > 0$. Production facilities are associated with a cost of $f_n(n)$, where $f_n(0) = 0$ and $df_n/dn,\ d^2 f_n/(dn)^2 > 0$. Henceforth, I refer to the sum of all of an entrepreneur's production facilities as that entrepreneur's firm. Once an entrepreneur starts operating in the market, all incurred costs are assumed to be sunk.

Each production facility of an entrepreneur is associated with a specific task, t, so that production facilities are uniformly distributed over tasks. Tasks are randomly and uniformly reassigned according to a Poisson process at rate $\delta > 0$. Low-skill workers who are able to perform the specific task have a productivity of $p(s, a) > p_{inf}$, where $\partial p/\partial s$, $\partial p/\partial a$, $\partial^2 p/\partial s \partial a > 0$ for all admissible values of managerial skills, s, and technology, a. Low-skill workers not able to perform the required task have a productivity of p_{inf}. I assume an analogous production technology for high-skill workers, except that their productivity exceeds that of the low-skill workers by a factor of $\theta_H > \theta_L = 1$, where the productivity of the low-skill workers, θ_L, is normalized to unity. Therefore, the production function

[12] All costs are modeled as perpetuities and all distributions are assumed to be continuously differentiable.

[13] For the sake of simplicity, I assume $d\Phi/df_a(0) > 0$, so that there always are some entrepreneurs willing to adopt the new technology at any skill level.

[14] I assume that there are no (binding) capacity constraints at production facilities, i.e., as many workers as desired may be employed at a single production facility.

of a firm is

$$y(s,a,m) = \int_0^m \tilde{\theta}(x)\tilde{p}(s,a,x)dx, \tag{1}$$

where m is the mass of workers employed at the respective firm, and $\tilde{p}(s,a,x) = p(s,a)$ if the worker x is able to perform the task required at the production facility, but p_{inf} otherwise. The function $\tilde{\theta}(\cdot)$ assumes the value θ_H for high-skill workers and is equal to unity for low-skill workers. All in all, high- and low-skill workers differ both in the share of tasks they are able to perform, α, and in productivity conditional on being able to perform a task, θ. Differences in productivity between worker groups within firms, θ, are not essential for the paper's main results on the impact of the technology diffusion process on the skill premium. However, modeling changes in high-skill workers' productivity, θ_H, over time allows me to assess the relative quantitative importance of technical progress complementing high-skill workers, on the one hand, and of skill-neutral technical progress and the technology diffusion process, on the other hand, for explaining the rise in the skill premium.

Type i unemployed workers, $i \in \{L, H\}$, receive an income flow of $\theta_i b$, which they have to forgo upon finding a job. Workers and entrepreneurs live forever. The time preference rate is denoted by ρ.

3.2. Matching and Wage Bargaining

Production facilities and workers are brought together pairwise through a sequential, random, and time-consuming search process. Let $\Gamma(\cdot)$ denote the equilibrium distribution of firms' $p(s,a)$-values, $n(p)$ the equilibrium mass of production facilities at a type p firm, and N the overall equilibrium mass of production facilities, i.e., $N = \tilde{N}\int_{p_{min}}^{p_{max}} n(x)d\Gamma(x)$, where p_{min} and p_{max} are the active firms' minimal and maximal p-values, respectively. Specifically, I assume workers to be contacted according to a Poisson process at rate $\lambda_w = N/M$. Since search and matching take place at the production facility level, a firm's contact rate depends positively on its mass of production facilities, i.e., a type p firm's contact rate is given by $n(p)$. The probability that an offer originates from a firm of type p or a lower type is given by

$$F(p) = \int_{p_{min}}^{p} n(x)d\Gamma(x) \bigg/ \int_{p_{min}}^{p_{max}} n(x)d\Gamma(x)\ . \tag{2}$$

Henceforth, I refer to $F(\cdot)$ as the workers' sampling distribution over firm types. Let $f(\cdot)$ designate the density function associated with $F(\cdot)$. The probability that a worker contacting a firm is a type i worker is simply given by the population share, i.e., M_i/M for $i \in \{L, H\}$.

Wages are bargained for by workers and employers in a complete information context. In particular, all agents who are brought to interact by the random matching process are perfectly aware of one another's types. All wage and job offers are also perfectly observable and verifiable. Wage contracts stipulate a fixed wage that can be renegotiated by mutual agreement only. Thus, renegotiations occur only if one party can credibly threaten the other to leave the match for good if the latter refuses to renegotiate. There are no renegotiation costs. Specifically, the wage is determined as the outcome of a Rubinstein (1982) infinite-horizon game of alternating offers, the precise structure and solution of which are characterized in Cahuc et al. (2006). This game delivers the generalized Nash bargaining solution, where the worker receives a constant share β of the match rent. The parameter β is referred to as the worker's bargaining power. In the remainder of this section I restrict the exposition to the case in which the worker is able to perform the task required at the production facility. A motivation is provided in Section 4.

Formally, let $V_i(w, p)$, $i \in \{L, H\}$, denote the lifetime utility of a low-skill and a high-skill worker, respectively, when employed at a type p production facility and paid a wage w. Two bargaining situations may arise in this framework: wage negotiations between an unemployed worker and an employer, e.g., a type p' production facility, and wage renegotiations that arise when employed workers are able to trigger competition between two employers, e.g., production facilities of types p and $p' > p$, for their services. Loosely speaking, the key difference between the two bargaining situations is the worker's fallback option. In the first case, the worker's fallback option is unemployment, which coincides with the lifetime utility of a worker who is employed at a production facility of type b and paid a wage $\theta_i b$, i.e., $V_i(\theta_i b, b)$. In the second case, the worker's fallback option is being employed at the less productive type p production facility while obtaining the entire match surplus, i.e., $V_i(\theta_i p, p)$. The outcome of the second bargaining game is the wage $\omega_i(p, p')$ at the type p' production facility, which leaves the worker with a value of $V_i(\theta_i p, p)$, the outside option, plus a share β of the match surplus $V_i(\theta_i p', p') - V_i(\theta_i p, p)$, i.e., $\omega_i(p, p')$ satisfies

$$V_i(\omega_i(p, p'), p') = V_i(\theta_i p, p) + \beta \left[V_i(\theta_i p', p') - V_i(\theta_i p, p) \right], \quad p' > p. \tag{3}$$

For $p = b$, this equation describes the negotiation outcome between an unemployed worker and a production facility of type p'.

Renegotiation takes place only if it is in the worker's interest. In particular, there exists a threshold $q_i(w, p)$ (formally defined by $\omega_i(q_i, p) = w$), such that *(i)* if $p' \leq q_i(w,p)$, then the worker keeps the current wage contract w at the type p production facility, *(ii)* if $p \geq p' > q_i(w,p)$, the worker obtains a wage raise $\omega_i(p',p) - w$ from the current employer, and *(iii)* if $p' > p$, the worker moves to the type p' production facility for a wage $\omega_i(p,p')$. Note that whenever $p' > p$, the wage $\omega_i(p,p')$ obtained at the new production facility can be smaller than the wage w paid in the previous job, because the worker expects higher wage raises at a production facility with a higher productivity. This option value effect implies that workers may be willing to take wage cuts just to move from low- to high-productivity production facilities.

The kind of alternating-offers infinite-horizon bargaining game à la Rubinstein that Cahuc et al. (2006) invoke as a foundation for the surplus splitting rule (3) predicts that as the breakdown rate of ongoing negotiations becomes large compared to the transition rates and the players' discount rates, the bargaining power is reduced to a function of the parties' relative response times only. Specifically, β is an increasing function of a worker's ability to formulate offers quickly (relative to the employer) and is otherwise independent of the arrival rate of job offers or any other structural parameter. So β can be considered as a separate structural parameter that specifically reflects the worker's ability to voice claims during bilateral negotiations with employers.

4. Equilibrium Characterization

In the following, I restrict the analysis to a subset of the equilibria that may arise in this environment. Specifically, I assume that p_{inf} is small enough so that the match surplus is negative if the task required by the production facility is not part of the tasks the worker is able to perform. This assumption simply rules out equilibria where matches are formed between workers and production facilities that do not match in tasks. Furthermore, note that as the breakdown rate of ongoing negotiations becomes large compared to the transition rates and the players' discount rates, workers and production facilities that do not match in tasks are deprived of the possibility of delaying agreement in anticipation that the match surplus may turn positive at a future point in time. Therefore, it suffices to restrict the analysis to immediate trade agreements and to disregard cases of continued bargaining. And finally, I only consider equilibria that arise as ρ tends to zero. Therefore, I assume that entrepreneurs maximize steady state profits as, for instance, in Burdett and Mortensen (1998). The advantage of this assumption is that it allows a concise representation of the various cutoff values.

Under these assumptions, the only matches formed are those in which the worker is able to perform the task required at the respective production facility. Similarly, if the task required at a production facility changes and the worker is not able to perform the new task, the match is dissolved. The resulting model is isomorphic in terms of worker flows to the canonical on-the-job search model, where the type i workers' offer-arrival rate and job destruction rate are given by $\lambda_i = \alpha_i \lambda_w$ and $\delta_i = (1 - \alpha_i)\delta$, respectively.

While the model is isomorphic to the canonical on-the-job search model, it provides a microfoundation for the observed differences in mobility patterns between skill groups. High- and low-skill workers compete for the same jobs and receive job offers at the same rate λ_w. However, low-skill workers are more likely to have to reject a job offer. They are only able to perform a smaller share of tasks and, therefore, are less likely to satisfy the task requirement of a job offer. Since only job offers with suitable task requirements are eventually of value to the worker, the effective high-skill workers' job-offer arrival rate, $\lambda_H = \alpha_H \lambda_w$, exceeds the low-skill workers' one, $\lambda_L = \alpha_L \lambda_w$. Similarly, low- and high-skill workers are exposed to changing task requirements at the same rate δ. However, low-skill workers are only able to perform a smaller share of tasks. Therefore, they are less likely to adapt to changing task requirements at the production facility and more likely to separate into unemployment. The resulting low-skill workers' job separation rate into unemployment, $\delta_L = (1-\alpha_L)\delta$, exceeds the high-skill workers' one, $\delta_H = (1 - \alpha_H)\delta$.

The proposed microfoundation for the differences in mobility between the skill groups is consistent with the empirical evidence. Conditional on a change in employment status, high-skill workers are more likely to switch employers since they receive more *suitable* job offers, and are less likely to separate into unemployment since they are more likely to adapt to changing task requirements. At the same time, this involves a higher likelihood on their part of switching tasks, since they are more likely to switch tasks instead of separating into unemployment. This is in line with the statistics presented in Section 2 on employer–employer transitions, separations into unemployment, and occupational changes. Furthermore, note that it represents a parsimonious rationalization of mobility patterns since the differences between skill groups in their employer–employer transitions, separations into unemployment, and occupational changes, are all ascribed to heterogeneity in one single parameter, i.e., the share of suitable tasks α.

Finally, under these assumptions, the type of a firm, $p(s, a)$, is closely related to both type i workers' marginal and average productivity, $\theta_i p$, at the respective firm. Therefore, I henceforth index firms by productivity p. I briefly characterize the equilibrium of the model in the remainder of this section. Detailed derivations are in Appendix B.

4.1. Workers

The competition of two employers of productivities q and $p > q$ over a worker's services yields the wage

$$\omega_i(q,p) = \theta_i q + \beta\theta_i(p-q) - (1-\beta)^2 \int_{\max\{q,p_{min}\}}^{p} \frac{\theta_i \lambda_i \bar{F}(x)}{\rho + \delta_i + \lambda_i \beta \bar{F}(x)} dx \qquad (4)$$

at the more productive employer p, where $\bar{F}(\cdot) = 1 - F(\cdot)$ designates the survivor function associated with $F(\cdot)$. The worker obtains the entire production flow that arises under the less efficient match, $\theta_i q$, plus a share β of the production flow surplus, $\theta_i(p-q)$, that is generated by the more efficient match, minus the option value (the last term on the right-hand side of the equation) that reflects the expected wage raises from future renegotiations. The wage agreement between an employer of productivity p and an unemployed worker is given by $\omega_i(b,p)$, since an unemployed worker's outside option corresponds to being employed at a production facility of productivity b. Therefore, only employers with a productivity of at least b are able to attract workers, and hence may have a non-zero steady state workforce.

Equation (4) shows that the wages of type i workers are solely a function of the competing employers' productivities. Intuitively, there are two factors that determine a worker's wage. On the one hand, there is an allocation effect. Being matched with a more productive employer generates a higher production flow. Bargaining then ensures that the worker enjoys a higher lifetime utility than would otherwise arise under a less efficient match. On the other hand, there is a rent share effect. Each worker that is employed at a facility of productivity p generates a production flow of $\theta_i p$ and, therefore, a flow rent of $\theta_i(p-b)$ relative to unemployment. For a given employer–employee match, the share of the rent $\theta_i(p-b)$ that the worker is able to appropriate is larger if the outside option is higher, i.e., if the productivity q of the other competing employer is higher. Therefore, workers benefit from competition between employers twice: from the induced reallocation from less to more productive production facilities and by being able to exploit other employers as outside options in the wage negotiations.

Since, by Equation (4), wages are solely a function of competing firms' productivities, it suffices to derive the distribution of workers over pq-pairs to obtain the wage distribution. That is, information about the productivity of the worker's current employer and the productivity of the most recent valuable outside option is sufficient. I use the Fokker–Planck formalism to derive the law of motion for the aggregate distribution that is consistent with individuals' laws of motion (see Bayer and Wälde, 2011). Detailed derivations are in Appendix B. It then can be shown that that the average wage of a type i worker at a firm of productivity p, $\bar{w}_i(p)$, satisfies

$$\bar{w}_i(p) = \theta_i\Bigg(p - (1-\beta)(\delta_i + \lambda_i\bar{F}(p))^2 \int_{p_{min}}^{p} \frac{\rho + \delta_i + \lambda_i\bar{F}(q)}{(\rho + \delta_i + \lambda_i\beta\bar{F}(q))(\delta_i + \lambda_i\bar{F}(q))^2} dq \\ - (1-\beta)(p_{min} - b)\frac{(\delta_i + \lambda_i\bar{F}(p))^2}{(\delta_i + \lambda_i)^2}\Bigg) \quad (5)$$

in the steady state equilibrium.

A worker's decision to become a high-skill worker takes a standard form. There is a cutoff value, denoted f_g^*, that equates the costs of becoming a high-skill worker to the expected income gain, i.e.,

$$f_g^* = (1-u_H)\bar{w}_H + u_H\theta_H b - (1-u_L)\bar{w}_L - u_L\theta_L b, \quad (6)$$

where $\bar{w}_i = \int_{p_{min}}^{p_{max}} \bar{w}_i(p)dJ_i(p)$ denotes the average wage rate and $u_i = \delta_i/(\delta_i + \lambda_i)$ denotes the unemployment rate of the type i workers, and $J_i(\cdot)$ designates the equilibrium distribution of type i workers over firm-productivity classes, i.e., $J_i(p) = \delta_i F(p)\Big/(\delta_i + \lambda_i\bar{F}(p))$. All workers who face an education cost of less than f_g^* incur the cost and become high-skill workers. All other workers remain low-skill workers. Therefore, the share of high-skill workers is given by $\Psi(f_g^*)$.

Equation (6) shows that there are two incentives to become a high-skill worker, i.e., a lower unemployment rate and higher wages conditional on being employed. The lower unemployment rate among high-skill workers is the result of a higher exit rate while being unemployed, $\lambda_H > \lambda_L$, and a higher labor market attachment once employed, $\delta_H < \delta_L$. Both channels reduce the unemployment rate and are both driven by the capability of performing a wider range of tasks. Indeed, rearranging the expression for the unemployment rate, i.e., $u_i = 1/(1 + \lambda_i/\delta_i)$, reveals that the unemployment rate is decreasing in the job-finding to separation rate ratio, i.e., λ_i/δ_i.

The skill premium is, intuitively, the result of three effects. First, there is a within-firm productivity effect. High-skill workers are more productive than low-skill workers at any firm. Specifically, high-skill workers productivity exceeds that of low-skill workers by a factor of θ_H. Rent sharing, as induced by the bargaining game (3), allows high-skill workers to profit from the higher production output. This is reflected in higher relative wages.

Second, there is an allocation effect. High-skill workers have a more favorable distribution over firm-productivity classes since their higher effective job-finding rate, $\lambda_H > \lambda_L$, and the lower separation rate into unemployment, $\delta_H < \delta_L$, foster the allocation from less to more productive firms. Rearranging the expression for the distribution of workers over firm-productivity classes, i.e., $J_i(p) = F(p) \Big/ (1 + (\lambda_i/\delta_i)\bar{F}(p))$, shows that the share of workers employed at firms of a productivity exceeding p is increasing in the job-finding to separation rate ratio, i.e., λ_i/δ_i, for any value of p.[15] In other words, a higher job-finding to separation rate ratio induces first-order stochastic dominance of the respective distribution. It follows immediately that the average production flow per worker, i.e., $\theta_i \int_{p_{min}}^{p_{max}} p dJ_i(p)$, is increasing in the job-finding to separation rate ratio as well. Rent sharing, as induced by the bargaining game (3), allows high-skill workers to profit from the more efficient allocation and the higher average match productivity.

Finally, there is a rent share effect. The higher effective job-finding rate, $\lambda_H > \lambda_L$, and the lower separation rate into unemployment, $\delta_H < \delta_L$, allow high-skill workers to appropriate a larger share of the production output for a given allocation of workers over firm-productivity classes. First, note that it follows from Equation (4) that the negotiated wage is increasing in the worker's outside option as given by the productivity q of the other competing employer. Furthermore, note that the average number of outside contacts that an employed worker can expect before the next unemployment period is simply given by the job-finding to separation rate ratio, i.e., λ_i/δ_i. While some contacts leave the employment relation unaffected and some contacts lead to employer–employer transitions, a share of the contacts leads to wage renegotiations allowing the worker to obtain wage raises from the current employer. High-skill workers experience more outside contacts per employment spell. Therefore, they are able to appropriate a larger rent share even for a given allocation of workers over firm-productivity classes.

[15] Cahuc et al.'s (2006) model can be interpreted as a productivity-ladder model. Workers climb the productivity-ladder by finding more productive firms, and fall down the productivity-ladder if they are forced to separate into unemployment. The job-finding to separation rate ratio, i.e. λ_i/δ_i, is the key determinant of a worker's expected position on the productivity-ladder.

4.2. Entrepreneurs

The type i workforce in a firm of productivity p with n production facilities, denoted by $m_i(n, p, \tau)$, evolves according to

$$\frac{dm_i(n, p, \tau)}{d\tau} = -(\delta_i + \lambda_i \bar{F}(p))m_i(n, p, \tau) + \alpha_i(M_i/M)n\,(u_i + J_i(p)(1 - u_i))\,. \quad (7)$$

Firms with a workforce of mass m_i, productivity p, and n production facilities lose workers when they separate into unemployment, $\delta_i m_i$, or are poached by more productive firms, $\lambda_i \bar{F}(p)m_i$. Firms attract workers who are unemployed, $\alpha_i(M_i/M)u_i n$, or poach workers from less productive firms, $\alpha_i(M_i/M)J_i(p)(1-u_i)n$. Therefore, the type i steady state workforce in a firm of productivity p and n production facilities is given by $m_i(p, n) = \delta_i \alpha_i(M_i/M)n \big/ (\delta_i + \lambda_i \bar{F}(p))^2$.

The steady state profit flow of a firm of productivity p with n production facilities, $\pi(p, n)$, is equal to the product of its average match rent and its workforce size:

$$\pi(p, n) = \sum_{i=L,H} (\theta_i p - \bar{w}_i(p))\, m_i(p, n). \quad (8)$$

Optimality requires that the marginal returns from additional production facilities equal marginal costs:

$$\partial\pi/\partial n\ (p, n) = df_n/dn\ (n). \quad (9)$$

The technology adoption decision takes a standard form. There exists a cutoff adoption cost, denoted $f_a^*(s)$, for each type of entrepreneur that equates the costs of adoption to the additional profits:

$$f_a^*(s) = (\pi(p(s, a_n)) - f_n(n(p(s, a_n)))) - (\pi(p(s, a_0)) - f_n(n(p(s, a_0))))\,, \quad (10)$$

where $\pi(p)$ denotes the equilibrium profits of a firm of productivity p. Entrepreneurs with lower adoption costs upgrade to the more advanced technology and entrepreneurs that have drawn a higher cost do not adopt the new technology. Let $s(p, a)$ denote the managerial skills necessary to achieve a productivity of p with the technology a. The share of firms with a productivity of less than p, $\Gamma(p)$, is equal to the share of entrepreneurs with managerial skills of at most $s(p, a_o)$ who operate the old technology and the share of entrepreneurs with managerial skills of at most $s(p, a_n)$ who adopt the new technology:

$$\Gamma(p) = \int_{\underline{s}}^{s(p,a_o)} (1 - \Phi\,(f_a^*(s)))\, d\tilde{\Gamma}(s) + \int_{\underline{s}}^{s(p,a_n)} \Phi\,(f_a^*(s))\, d\tilde{\Gamma}(s). \quad (11)$$

The decision to participate in the market and to produce takes a standard form as in the previous cases. However, it yields a managerial skill cutoff rather than a cost cutoff. Specifically, the least skilled entrepreneur producing in the market operates the new technology with a skill level satisfying the condition that the resulting profits equal the sum of the participation and technology adoption costs:

$$\pi(p(s_{min}, a_n)) - f_n(n(p(s_{min}, a_n))) = f_p + \underline{f_a}, \tag{12}$$

where $\underline{f_a}$ denotes the lowest possible draw of an adoption cost and is assumed to be zero. Furthermore, it follows that the least productive firm producing in the market has a productivity of $p(s_{min}, a_n)$.

The mass of entrepreneurs, $\tilde{N}$, entering the market is determined by a free entry condition. The expected profits from entering the market equal the entry cost:

$$\begin{aligned} f_e = \int_{p_{min}}^{p_{max}} \pi(p) d\Gamma(p) - \int_{p_{min}}^{p_{max}} f_n(n(p)) d\Gamma(p) \\ - \int_{\underline{s}}^{\bar{s}} \int_0^{f_a^*(s)} f_a d\Phi(f_a) d\tilde{\Gamma}(s) - f_p \left[1 - \Gamma(p_{min})\right], \end{aligned} \tag{13}$$

where the first term on the right-hand side reflects the expected profit flow, the second term the production facility setup costs, the third term the technology adoption costs, and the last term the participation costs.

5. Quantitative Exercise

In this section I assess the quantitative implications of the model. A numerical solution of the model necessitates functional form assumptions. I assume a firm's productivity to be given by the product of the entrepreneur's skill-level and the technology, i.e., $p(s, a) = sa$. The production facility setup costs are modeled as a power function, i.e., $f_n(n) = \bar{\eta} n^{\eta} / \eta$ for $\eta > 2$ and $\bar{\eta} > 0$. All distributions are assumed to belong to the generalized Pareto family, i.e.,

$$\Upsilon(x; \xi, \mu, \sigma) = \begin{cases} 1 - (1 + \xi \frac{x - \mu}{\sigma})^{-1/\xi} & \text{for } \xi \neq 0 \\ 1 - e^{-\frac{x-\mu}{\sigma}} & \text{for } \xi = 0 \end{cases}, \tag{14}$$

where $\xi \in (-\infty, \infty)$, $\mu \in (-\infty, \infty)$, $\sigma \in (0, \infty)$ and $x \geq \mu$. Henceforth, I assume ξ, σ, and μ to be equal to 0, 1, and 0, respectively, if not otherwise stated.

The equilibrium of the model is characterized by three algebraic equations and two differential equations. The algebraic equations (6), (12), and (13) determine the share of high-skill workers, the lower productivity cutoff, and the mass of entrants, respectively. The differential equations arise from the entrepreneurs' first order condition (9) and are

$$v(p) = \bar{\eta}\left(-\frac{ds/dp\ (p)}{\gamma(p)\tilde{N}/M}\right)^{\eta-1}, \quad \text{and} \tag{15}$$

$$dv/dp\ (p) = \sum_{i=L,H} \frac{\alpha_i\theta_i\delta_i(1-\beta)M_i/M}{(\delta_i + \beta\alpha_i s(p))(\delta_i + \alpha_i s(p))}, \tag{16}$$

where $s(p) = \lambda_w \bar{F}(p)$, and where $\gamma(\cdot)$ designates the density associated with the productivity distribution $\Gamma(\cdot)$ that is given by Equation (11). The boundary conditions are given by $v(p_{min}) = \sum_{i=L,H} \alpha_i\theta_i\delta_i(p_{min} - b)M_i/M \Big/ (\delta_i + \alpha_i N/M)^2$ and $s(p_{max}) = 0$. Note that the system of equations is not a system of ordinary differential equations. However, it can be represented as a higher dimensional system of ordinary and algebraic equations and, therefore, solved by standard numerical algorithms. The details are in Appendix C.

For the paper's quantitative part, I introduce an additional shock, i.e., the visibility shock χ. The shock χ affects the contact rate of the firm, i.e., $\chi n(p)$, and loosens the otherwise one-to-one mapping between firm size and productivity. Clearly, firm size is the outcome of various factors and random events. Summarizing all those factors and events in one single shock is a pragmatic reduced form approach, yet, prima facie, not inappropriate in this context. Finally, I assume that the χ-shock has a mean of unity and is realized after the firm starts producing in the market. Therefore, entrepreneurs' expectations remain unaffected, as do the respective optimality conditions. However, the visibility shock does affect the size distribution of firms and the skill composition within firm size classes in equilibrium.

5.1. Calibration

In this section I calibrate the model. First, I calibrate the transition parameters using monthly transition statistics from the year 1997. Loosely speaking, the calibration of the transition parameters is to be regarded as a separate exercise and is independent of the calibration of all the other parameters. Second, I calibrate all the other parameters, targeting in particular the productivity distribution in 1977 and 1997, the size distribution in 1977, and the supply of skill in 1977 and 1997.

I use a simple unconditional calibration strategy for the transition parameters λ_L, λ_H, δ_L, and δ_H. The details are in Appendix C. I then impose the same underlying job-offer arrival rate, λ_w, and task-switch rate, δ, for both groups, and choose values for α_H and α_L to rationalize the differences in the parameters λ_L, λ_H, δ_L, and δ_H. This reflects the model assumption that low- and high-skill workers compete for the same jobs and are exposed to the same informational frictions. All differences in mobility patterns between the skill groups are attributed to differences in the share of tasks the workers are able to perform, i.e., α.[16]

Table 1 provides transition statistics for U.S. manufacturing in 1997. After one month, on average 94.6 percent of production workers and 96.3 percent of non-production workers are still working for the same employer. This is above the economy-wide average of 93.0 percent. On the other hand, only 2.1 percent of production workers and 2.0 percent of non-production workers switch employers, which is below the economy-wide average of 3.0 percent. Separations into unemployment are also below the economy-wide average of 4.0 percent, at 3.3 percent for production and at 1.7 percent for non-production workers.

Table 1 shows the calibration of the transition parameters as well. All estimates are per month. The average production workers' employment spell before the next unemployment spell, $1/\delta_L$, is 2.4 years. The average non-production workers' employment spell, $1/\delta_H$, is 4.6 years. Conditional on staying with the same employer, both production and non-production workers face a task-switch, $(1/\delta)$, on average, every 11 months. The average number of outside contacts that an employed worker can expect before the next unemployment period, λ_w/δ_i, is 3.0 for production workers and 5.7 for non-production workers. However, only 62.7 percent of outside contacts have suitable task requirements in the case of production workers, and 80.4 percent in the case of non-production workers (α_i). Therefore, the effective number of outside contacts, λ_i/δ_i, amounts to 1.9 and 4.6 among production and non-production workers, respectively. It is primarily this difference between the two groups that is exploited in the subsequent quantitative analysis. However, the difference in mobility according to this measure, λ_i/δ_i, is substantial. Non-production workers expect more than twice as many suitable outside contacts per employment spell than production workers.[17]

[16] Alternatively, one may depart the from the assumptions of identical λ_w and δ parameters and, for instance, exploit information on occupational mobility to identify the additional parameters. However, it proves challenging to discipline the model with the given data sets and I leave this extension to further research.

[17] Note that the estimates are based on all worker groups without any restrictions, e.g., including

	production workers	non-production workers
statistics		
same employer	0.9457	0.9628
new employer	0.0214	0.0200
unemployment	0.0328	0.0172
parameters		
job-offer arrival rate (λ_w)	0.1032	0.1032
effect. job-offer arrival rate (λ_i)	0.0648	0.0830
task-switch rate (δ)	0.0924	0.0924
job-destruction rate (δ_i)	0.0345	0.0181
worker mobility (λ_i/δ_i)	1.8795	4.5916
versatility (α_i)	0.6273	0.8044

Table 1: Transition statistics and parameters for 1997. Statistics summarize monthly changes (employment shares) in employment status of workers employed in U.S. manufacturing. Parameters obtained from restricted model imposing identical λ_w and δ values for both production and non-production workers. Author's calculations based on the Current Population Survey Basic Monthly data as provided by the National Bureau of Economic Research (`http://www.nber.org/data/cps_basic.html`). See Appendix A and Appendix C for further details.

Table 2 provides an overview of the calibration of all the other model parameters. I set the new technology parameter, a_n, to 1.32, so that the annualized productivity growth rate among firms that adopt the new technology is 1.4 percent. The distribution of managerial skills, $\sigma_s = 1.33$, and the distribution of the technology adoption costs, $\sigma_a = 390$, are calibrated to match the standard deviation of the log-revenue-productivity in U.S. manufacturing of 0.45 in 1977 and 0.49 in 1997. The productivity dispersion estimates are from Hsieh and Klenow (2009). First, note that the estimates are for revenue-productivity dispersion, i.e., the product of physical productivity and a firm's output price. I use a broad notion of productivity dispersion, allowing information and communication technologies to affect firms in various ways. Second, the estimates are for total factor productiv-

both part-time and full-time workers and all demographics. Furthermore, note that separations out of employment encompass both unemployment and movements out of the labor force.

parameter	target	target value	parameter value
a_n	productivity growth rate	1.4%	1.32
σ_s	log-revenue-productivity s.d. in 1977 (Hsieh and Klenow, 2009)	0.45	1.33
σ_a	log-revenue-productivity s.d. in 1997 (Hsieh and Klenow, 2009)	0.49	390
$\bar{\eta}$	job-offer arrival rate in 1997	0.103	0.29
η	employment share of small establishments in 1977	0.59	3.50
f_e	worker–establishment ratio in 1977	52.79	30.53
f_p	establishment exit rate in 1977 (Business Dynamics Statistics)	10%	0.34
ξ_χ	production worker share at small establishments in 1977	0.76	0.71
ξ_g	share of prod. workers in 1977	0.74	12.71
σ_g	share of prod. workers in 1997	0.72	0.45
$\theta_{H,1977}$	skill premium in 1977	1.53	1.33
$\theta_{H,1997}$	skill premium in 1997	1.74	1.44
β	Cahuc et al. (2006)	-	0.05

Table 2: Calibration of model parameters. Distributional parameters ξ, σ and μ are set to 0, 1 and 0, respectively, if not otherwise stated. Transition parameters are displayed in Table 1.

ity and not labor productivity. The productivity estimates are, therefore, intended to control for a rising inter-industry dispersion of capital–labor ratios, which is at the core of the mechanism proposed by Caselli (1999).

The facility setup cost parameter $\bar{\eta}$ equals 0.29, so that the job-offer arrival rate, λ_w, is 0.103 in 1997. The curvature parameter of the facility setup costs, η, and the distribution of the visibility shock, χ, are calibrated to match moments of the size distribution. Specifically, I set η to 3.5 to obtain an employment share of 59 percent of firms with less than 500 employees in 1977. The shape parameter of the visibility shock, ξ_χ, equals 0.71, resulting in a production worker share of 0.76 at firms with less than 500 employees, and the scale parameter, σ_χ, equals $1 - \xi_\chi$, so that the visibility shock has a mean of one. The entry cost, f_e, amounts to 30.53, resulting in a worker–firm ratio of 52.79 in 1977. The participation cost,

$f_p = 0.34$, is calibrated to obtain an exit rate of entrepreneurs of 10 percent in 1977, where the exit rate is defined as $\tilde{\Gamma}(p_{min}) - \tilde{\Gamma}(b) \big/ 1 - \tilde{\Gamma}(b)$.

The distribution of the education costs, $\xi_g = 12.71$ and $\sigma_g = 0.45$, is calibrated to match the shares of production workers of 0.74 in 1977 and 0.72 in 1997. Differences in mobility between production and non-production workers are not sufficient to explain the skill premium in 1977. Therefore, I assume non-production workers to be more productive. Specifically, I set the non-production workers' productivity parameter in 1977, $\theta_{H,1977}$, to 1.33 in order to obtain a skill premium of 1.53. Similarly, the rise in productivity dispersion across firms is not sufficient to explain the entire increase in the skill premium. Therefore, I assume non-production workers' productivity to increase by 8.5 percent from 1977 to 1997 in order to obtain a skill premium of 1.74 in 1997. In the subsequent analysis I decompose the skill premium and changes in the skill premium to quantify the contribution of the different factors. Lacking comparable estimates for the United States, I impose a rent share parameter, β, of 0.05, i.e., the average estimate from Cahuc et al. (2006) for French manufacturing.[18]

5.2. *Model Predictions*

In this section I review the quantitative implications of the model. The share of firms employing the new technology is predicted by the model to be 13 percent in 1997. Figure 4 depicts on the left-hand side the adoption rate by managerial skills. The technology adoption rate is increasing in the level of managerial skills. This pattern is induced by the assumption that the adoption of the new technology is associated with fixed costs. More skilled managers operate on average larger firms and, therefore, can spread the fixed costs over a larger base. Furthermore, note that the technology adoption rate is not zero for all managerial skill levels below 1.14, i.e., the productivity cutoff in 1977. Entrepreneurs with the respective managerial skill levels do not participate in the market in 1977. However, some of these entrepreneurs who have drawn low technology adoption costs adopt the new technology in 1997 and successfully participate in the market.

As targeted, the standard deviation of log-productivity, i.e., the logarithm of $p(s, a)$, is 0.45 in 1977 and 0.49 in 1997. Figure 4 depicts on the right-hand side the change in the scaled productivity distribution, i.e., the density of active firms'

[18]Cahuc et al. (2006) estimates reveal considerable heterogeneity across skill groups. Low-skill workers tend to have lower bargaining power. While this heterogeneity in bargaining power might provide another channel through which the skill premium is affected, it is outside this paper's scope and I leave it to further research.

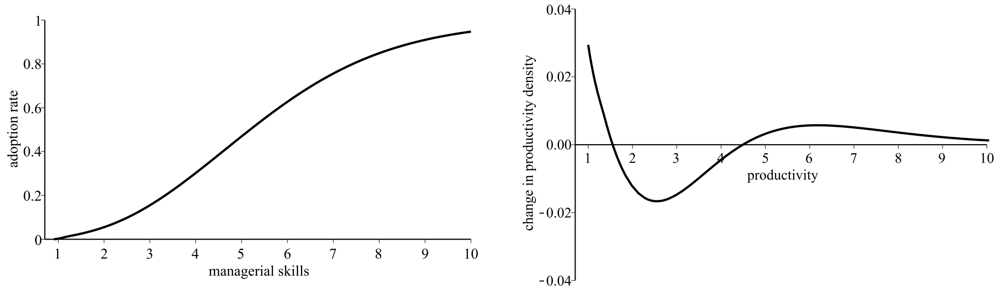

Figure 4: On the left-hand side: technology adoption rate by managerial skills (s) as predicted by the model. On the right-hand side: change in scaled productivity density (p) between 1977 and 1997 as predicted by the model. Model parameters as depicted in Tables 1 and 2.

$p(s,a)/p_{min}$-values, between 1977 and 1997. Since the productivity density integrates to one in each year, the change in the density across years must integrate to zero. Thus, the height at each (relative) productivity level measures the growth in the respective firms' share relative to the whole. Figure 4 shows that the availability of the new technology leads to a polarization of the productivity distribution. While the share of high- and low-productivity firms increases, the share of firms with intermediate productivity decreases.

Essentially, the positive correlation between the adoption rate and managerial skills leads to an increase in the share of high-productivity firms and a decrease in the share of low-productivity firms from 1977 to 1997, since it is predominantly the most productive firms that experience an increase in their level of productivity as the new technology becomes available. However, changes in the productivity distribution are driven by the extensive margin as well. On the one hand, as successful entrepreneurs adopt the more advanced technology, the least productive firms are pushed out of the market. The lower productivity cutoff increases from 1.14 to 1.20. On the other hand, the availability of the more advanced technology makes it possible for some low-skill entrepreneurs who would have otherwise remained inactive to participate in the market. All in all, the effects at the extensive margin increase the share of low-productivity firms.

Table 3 summarizes the model predictions and the data counterparts for small and large firms. As targeted, the model matches the employment share of 59 percent of establishments with less than 500 employees and the skill share of 24 percent at the respective establishments in the year 1977. In 1997 the observed employment share is 68 percent and the observed skill share is 27 percent. However, the model predicts the employment share of establishments with less than 500 employees to slightly fall to 58 percent. The skill share rises to 26 percent

	data			model			%
	1977	1997	Δ	1977	1997	Δ	
small firms' employment share	0.59	0.68	0.09	0.59	0.58	-0.01	-11
skill share at small firms	0.24	0.27	0.03	0.24	0.26	0.02	66
value added ratio	1.38	1.73	0.35	1.33	1.45	0.12	34
Δsize premium	-0.28	-0.15	0.13	0.24	0.38	0.13	100

Table 3: Statistics for small and large firms as observed in the data and predicted by the model. Firms with less (more) than 500 employees defined as small (large). Values in the first column from the right are changes predicted by the model (column 7) in percent of the data counterparts (column 4). Model parameters as displayed in Tables 1 and 2. Author's calculations based on Census of Manufactures reports. See Appendix A and Appendix C for further details.

in the model. With the given calibration the model does not attribute the sharp increase in the employment share of small establishments to the advent of information and communication technologies. Other factors are likely to have influenced the size distribution over time. Furthermore, some aspects of information technologies, that are not modeled here, may have induced firms to reduce their size (see, e.g., Brynjolfsson et al., 1994).

The model predicts the value added per worker at establishments with at least 500 employees to be 1.33 times higher than at establishments with less than 500 employees in 1977. The corresponding ratio in the data is 1.38. While only the revenue-productivity dispersion is targeted in 1977, the model also matches the difference in value added per worker between small and large firms in 1977 fairly well. However, the rise in revenue-productivity dispersion from 0.45 in 1977 to 0.49 in 1997 is not sufficient to generate a rise in the value added ratio of the same magnitude in the model as in the data. The value added ratio increases from 1.38 in 1977 to 1.73 in 1997 in the data. On the other hand, the ratio increases form 1.33 in 1977 to only 1.45 in 1997 in the model.

The firm size wage premium for production workers in the model is 37 percent in 1977 compared to 45 percent in the data. The firm size wage premium for non-production workers is 61 percent in 1977 compared to 17 percent in the data. The differential firm size wage premium between non-production and production workers increases by 13 percentage points in the model and in the data as well. The observed skill premium is around 53 percent in 1977. It increases by 21 percentage points to 74 percent in 1997. As targeted, the model generates coinciding skill premium patterns. In summary, the model captures 100 percent of the change in the skill premium and 100 percent of the change in the differential firm size wage premium.

The job-finding rate, λ_w, rises in the model from 0.080 in 1977 to 0.103 in 1997. The change in the job-finding rate is associated with an increase by 19.5 percent in the average unconditional employer–employer transition rate and by 8.8 percent in the average unconditional rate at which workers switch tasks.[19] This is broadly consistent with the increase in annual occupational mobility at the three-digit level over the same time period by 25 percent as reported by Kambourov and Manovskii (2008).[20]

The model also predicts that workers in larger firms switch occupations more often inside the firm and are less likely to separate from the firm. This is well in line with stylized facts stressed in the literature (see, e.g., Idson, 1989; Papageorgiou, 2010). The higher rate of occupational changes at large firms results from a higher share of versatile workers at large firms in the model. Versatile workers exhibit a higher degree of inter-firm mobility. Therefore, they are more likely to match with high-productivity and at the same time large firms. Since versatile workers are more likely to adapt to changing task requirements, rather than to separate into unemployment, the frequency of occupational changes at large and productive firms exceeds that of small firms. The higher share of versatile workers at large firms also leads to a lower separation rate, since this group is less likely to separate into unemployment. Furthermore, workers at larger and more productive firms are less likely to receive job offers that induce them to switch employers. Indeed, the employer–employer transition rate, $\lambda_i\bar{F}(p)$, is decreasing in the productivity of the firm.[21]

[19]The analytical expression for type i workers' unconditional employer–employer transition rate is $(\delta_i/\lambda_i)(\delta_i+\lambda_i)ln\,(\delta_i+\lambda_i/\delta_i)-\delta_i$, and the rate for task-switches is $(\delta_i/\lambda_i)(\delta_i+\lambda_i)ln\,(\delta_i+\lambda_i/\delta_i)-(1-\alpha_i)\delta_i$. See Appendix C.2 for further details.

[20]Moscarini and Thomsson (2007), relying on the Current Population Survey data for the years 1994 and onwards, find that 67 percent of job-to-job movers change occupations as well.

[21]However, high-skill workers have a higher effective job-offer arrival rate, λ_i, which mitigates the differences between small and large establishments in the average employer–employer separation rate over all worker groups.

5.3. Counterfactual Decompositions

In this section I proceed with several counterfactual decompositions in order to gain insights into the factors determining the skill premium.[22] First, I briefly analyze the skill premium in 1977. I set high-skill workers' productivity in 1977, $\theta_{H,1977}$, to unity, so that low- and high-skill workers are equally productive at a specific firm. The counterfactual skill premium, i.e., $(1/\theta_{H,1977})\bar{w}_{H,1977}/\bar{w}_{L,1977}$, amounts to 15 percentage points. Therefore, differences in mobility between production workers and non-production workers are predicted to account for 28 percent of the skill premium of 53 percentage points in 1977.[23]

Next, I turn to changes in the skill premium between 1977 and 1997. I construct two counterfactual skill premia and differential firm size wage premia for the year 1997. To isolate the effects of skill-neutral technical progress and the technology diffusion process, I set high-skill workers' productivity in 1997, $\theta_{H,1997}$, to its 1977 value of 1.33. Hence, high-skill workers' counterfactual average wage in 1997 is $(\theta_{H,1977}/\theta_{H,1997})\bar{w}_{H,1997}$. In other words, I assume no additional improvements in the technology complementing high-skill workers. The results are displayed in Table 4 under "Skill-Neutral Technical Change." The model still explains the entire change in the differential firm size wage premium. About one-third of the rise in the skill premium is ascribed to skill-neutral technical change and the technology diffusion process. In summary, differences in inter-firm mobility between non-production and production workers explain 28 percent of the skill premium in 1977 and one-third of the increase in the skill premium from 1977 to 1997.

The heterogeneous impact of skill-neutral technical change on skill groups results from differences in inter-firm mobility between the skill groups and not technology–skill complementarity. Intuitively, it is the polarization of the productivity distribution as depicted in Figure 4 that contributes to the rise in the

[22]The following results are counterfactual decompositions and do not account for general equilibrium effects, i.e., only the direct effects of parameter changes are considered.

[23]My approach to exploit differences in mobility patterns between worker groups in order to explain wage patterns is related to the literature that exploits worker in- and outflows at the firm level in order to identify the elasticity of labor supply to the firm with respect to the wage. It is the supply elasticity *to the firm* that matters for firms' wage policies, not necessarily the supply elasticity *to the market*. For instance, Ransom and Oaxaca (2010) find that women supply labor less elastically to the firm than men in the United States retail grocery industry. Therefore, it is consistent with profit-maximizing discrimination against women. Quantitatively, they find that the estimated differences in supply elasticities explain reasonably well the lower relative pay of women.

	model			%
	1977	1997	Δ	
Baseline Specification				
skill premium	1.53	1.74	0.21	100
Δsize premium	0.24	0.38	0.13	100
Skill-Neutral Technical Change				
skill premium	1.53	1.60	0.07	34
Δsize premium	0.24	0.38	0.13	100
Skill-Biased Technical Change				
skill premium	1.53	1.66	0.13	63
Δsize premium	0.24	0.24	0.00	0

Table 4: Skill premium and differential firm size wage premium counterfactuals for U.S. manufacturing as implied by the model. Values in the first column from the right are predicted changes in percent of the actually observed changes. Model parameters as displayed in Tables 1 and 2. See Appendix A and Appendix C for further details.

returns to inter-firm mobility. The polarization of the productivity distribution translates into a polarization of workers' employment opportunities. Workers are more likely to obtain job-offers either form low- or high-productivity firms rather than firms with intermediate productivity. Hence, the pronounced importance of inter-firm mobility, since it enables workers to move from an unfavorable extreme to a favorable extreme, i.e., form low-productivity to high-productivity firms.

To isolate the effect of improvements in the technology complementing high-skill workers, I assume 1977's environment and only consider a change in the technology complementing high-skill workers, θ_H. Hence, high-skill workers' counterfactual average wage in 1997 is $(\theta_{H,1997}/\theta_{H,1977})\bar{w}_{H,1977}$. The results are displayed in Table 4 under "Skill-Biased Technical Change." The firm size wage premium remains unaffected by skill-biased technical change. Rising skill–technology complementarity increases high-skill workers' wages at both small and large firms, leaving the firm size wage premium unaffected. The model ascribes 63 percent of the increase in the skill premium to improvements in the technology complementing high-skill workers.

To provide further insights into the mechanisms driving the change in the skill premium, an additional counterfactual exercise follows. First, I divide the firm size distribution into four size categories, $j \in \{1, .., 4\}$: less than 100 employees ($j = 1$), less than 500 and at least 100 employees ($j = 2$), less than 1000 and at least 500 employees ($j = 3$), and at least 1000 employees ($j = 4$). Then I calculate the average wage in each size category and the employment share of each size category for both worker types and for the years 1977 and 1997. Let $\bar{w}_{i,j,t}$ and $d_{i,j,t}$ denote the average wage and employment share, respectively, of type i workers in the firm size class j in the year t. Consider following expression for the non-production and production workers wage ratio:

$$\bar{w}_H / \bar{w}_L \,(s_w, s_d, t) = \sum_{j=1..4} \bar{w}_{H,j,t} d_{H,j,t} \Bigg/ \sum_{j=1..4} \bar{w}_{s_w,j,t} d_{s_d,j,t} \,, \tag{17}$$

for $s_w, s_d \in \{L, H\}$ and $t \in \{1977, 1997\}$. For $s_w = s_d = L$ and $t = 1977$, the equation simply yields the skill premium in the year 1977. Other combinations of s_w and s_d yield counterfactual skill premia, which are discussed in the following.

The first counterfactual exercise consists in positing that low- and high-skill workers earn the same wage within a given firm size class. Hence, the difference in wages between skill groups is solely driven by the difference between their distributions over firm size classes. I refer to this difference in relative wages as the "allocation premium." Specifically, the respective counterfactual skill premium in the year t is then $\bar{w}_H / \bar{w}_L \,(s_w = H, s_d = L, t)$. Table 5 juxtaposes the allocation premium from the data with that predicted by the model. The allocation premium is only one percentage point in the data and increases by one percentage point from 1977 to 1997. This represents five percent of the overall increase in the skill premium. In the model, the allocation premium amounts to four percentage points and increases by three percentage points. This represents fourteen percent of the overall increase in the skill premium.

The second counterfactual exercise consists in assuming that both high-skill workers and low-kill workers have the same distribution over firm size classes. Hence, the differences in wages between the skill groups are solely driven by the differences between the skill groups in wages by firm size class. I refer to this difference in relative wages as the "wage premium." Specifically, the corresponding counterfactual skill premium in the year t is then $\bar{w}_H / \bar{w}_L \,(s_w = L, s_d = H, t)$. Table 5 juxtaposes the wage premium from the data with that predicted by the model. The wage premium is 49 percentage points in the data in 1977 and increases by 20 percentage points from 1977 to 1997. This represents 95 percent of the overall increase in the skill premium. The model predicts also a wage premium of 49 percentage points in 1977. It increases by 17 percentage points, which represents 81 percent of the overall increase in the skill premium.

	data				model			
	1977	1997	Δ	%	1977	1997	Δ	%
skill premium	1.53	1.74	0.21	-	1.53	1.74	0.21	-
allocation premium	1.01	1.02	0.01	5	1.04	1.07	0.03	14
wage premium	1.49	1.69	0.20	95	1.49	1.66	0.17	81
skill-neutral tech. change	-	-	-	-	1.49	1.53	0.04	24
skill-biased tech. change	-	-	-	-	1.49	1.62	0.13	76

Table 5: Skill premium counterfactuals for U.S. manufacturing and model counterparts. Values in the fourth and eighth column are the counterfactual changes in percent of the wage premium change (last two rows) and in percent of the skill premium change (preceding two rows). Author's calculations based on Census of Manufactures reports. Model parameters as displayed in Tables 1 and 2. See Appendix A and Appendix C for further details.

The final two counterfactual exercises decompose the change in the wage premium. Specifically, they illustrate the relative importance of skill-neutral and skill-biased technical progress. In the first exercise, I assume high-skill workers' productivity in 1997, $\theta_{H,1997}$, to equal its value in 1977. Hence, changes in the wage premium are solely driven by skill-neutral technical change and the technology diffusion process. Specifically, the corresponding counterfactual skill premium in the year 1997 is then $(\theta_{H,1977}/\theta_{H,1997})\bar{w}_H/\bar{w}_L\ (s_w = L, s_d = H, 1997)$. Table 5 displays the change in the wage premium due to skill-neutral technical progress. It is 24 percent of the overall increase in the wage premium. The second counterfactual exercise consists again in calculating the wage premium, however, I assume that wage changes are only due to an improvement in high-skill workers' productivity, θ_H. Specifically, the corresponding counterfactual skill premium in 1997 is then $(\theta_{H,1997}/\theta_{H,1977})\bar{w}_H/\bar{w}_L\ (s_w = L, s_d = H, 1977)$. Table 5 displays the wage premium due to skill-biased technical change. It accounts for 76 percent of the increase in the wage premium.

6. Conclusion

In this paper, I stress the importance of *(i)* versatility, i.e., the ability to perform a wide range of activities or tasks even across occupations, and *(ii)* heterogeneity in technology adoption across firms, for understanding the evolution of the skill premium over the last decades. Versatility enables workers to overcome structural impediments to employer–employer transitions and increases their inter-firm mobility. The returns to inter-firm mobility are higher when the disparities in profitability between firms are higher. Insofar as technical progress favors the more productive over less productive firms, the productivity dispersion across firms rises and, hence, so do the returns to inter-firm mobility. High-skill (i.e., versatile) workers obtain higher wage raises than do other workers.

In a quantitative analysis, I exploit the different inter-firm mobilities of the two skill groups and the rising productivity dispersion across firms in order to explain the rise in the skill premium from 1977 to 1997. The microstructure of the model is in line with the data, and the key patterns of the proposed link between productivity dispersion and the skill premium are observed in the data as well. In particular, the model postulates a close link between the skill premium and the differential firm size wage premium between high- and low-skill workers. All in all, the model ascribes one-third of the sharp increase in the skill premium in U.S. manufacturing from 1977 to 1997 to skill-neutral technical change and the technology diffusion process itself.[24]

While the model already generates substantial returns to versatility and attributes a sizable share of the increase in the skill premium to differences in versatility and the consequent differences in inter-firm mobility between skill groups, it does not capture all potentially quantitatively important channels. In particular, high-skill workers make up a disproportionately large share of the workforce in large and productive firms as a consequence of their versatility and, hence, their higher inter-firm mobility. This skill abundance at the most productive and large firms may favor the implementation of skill-biased technologies by these firms and further raise the returns to versatility.[25]

[24] Stijepic (2015c) studies the link between the returns to versatility and log-sales per worker dispersion across firms in the late 1990s. He finds a significantly higher versatility wage premium in industries with higher log-sales per worker dispersion across firms. Specifically, an increase in the standard deviation of sales by 0.5 is estimated to raise above-median versatile workers relative wage by 11 percentage points. For a subsample of industries he also analyzes the relation between changes in the college wage premium and changes in sales dispersion within industries between the late 1970s and 1990s. An increase in the standard deviation of log-sales per worker by 0.1 is estimated to increase the college wage premium by 15 percentage points.

[25] See Acemoglu (2002) who stresses the endogeneity of the direction and the bias of technical change.

References

Abowd, J. M., Kramarz, F., Margolis, D. N., 1999. High Wage Workers and High Wage Firms. Econometrica 67 (2), 251–334.

Acemoglu, D., 2002. Directed Technical Change. Review of Economic Studies 69 (4), 781–809.

Acemoglu, D., 2003. Patterns of Skill Premia. Review of Economic Studies 70 (2), 199–230.

Acemoglu, D., Autor, D., 2011. Skills, Tasks and Technologies: Implications for Employment and Earnings. Vol. 4 of Handbook of Labor Economics. Elsevier, Ch. 12, pp. 1043–1171.

Aghion, P., Howitt, P., Violante, G. L., 2002. General Purpose Technology and Wage Inequality. Journal of Economic Growth 7 (4), 315–45.

Alvarez, F., Shimer, R., 2011. Search and Rest Unemployment. Econometrica 79 (1), 75–122.

Andersson, F., Davis, E. E., Freedman, M. L., Lane, J. I., Mccall, B. P., Sandusky, K., 2012. Decomposing the Sources of Earnings Inequality: Assessing the Role of Reallocation. Industrial Relations: A Journal of Economy and Society 51 (4), 779–810.

Bayer, C., Wälde, K., 2011. Describing the Dynamics of Distributions in Search and Matching Models by Fokker-Planck Equations. Working Papers 1110, Gutenberg School of Management and Economics, Johannes Gutenberg-Universität Mainz.

Brynjolfsson, E., Malone, T. W., Gurbaxani, V., Kambil, A., 1994. Does Information Technology Lead to Smaller Firms? Management Science 40 (12), 1628–1644.

Burdett, K., Mortensen, D. T., 1998. Wage Differentials, Employer Size, and Unemployment. International Economic Review 39 (2), pp. 257–273.

Cahuc, P., Postel-Vinay, F., Robin, J.-M., 2006. Wage Bargaining with On-the-Job Search: Theory and Evidence. Econometrica 74 (2), 323–364.

Card, D., Heining, J., Kline, P., 2013. Workplace Heterogeneity and the Rise of West German Wage Inequality. The Quarterly Journal of Economics 128 (3), 967–1015.

Caselli, F., 1999. Technological Revolutions. American Economic Review 89 (1), 78–102.

Dunne, T., Foster, L., Haltiwanger, J., Troske, K. R., 2004. Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment. Journal of Labor Economics 22 (2), 397–430.

Entorf, H., Gollac, M., Kramarz, F., 1999. New Technologies, Wages, and Worker Selection. Journal of Labor Economics 17 (3), pp. 464–491.

Faggio, G., Salvanes, K., Reenen, J. V., 2007. The Evolution of Inequality in Productivity and Wages: Panel Data Evidence. CEP Discussion Papers dp0821, Centre for Economic Performance, LSE.

Faggio, G., Salvanes, K. G., Reenen, J. V., 2010. The Evolution of Inequality in Productivity and Wages: Panel Data Evidence. Industrial and Corporate Change 19 (6), 1919–1951.

Fallick, B., Fleischman, C. A., 2004. Employer-to-Employer Flows in the U.S. Labor Market: The Complete Picture of Gross Worker Flows. Finance and Economics Discussion Series 2004-34, Board of Governors of the Federal Reserve System (U.S.).

Gremillion, L. L., 1984. Organization Size and Information System Use: An Empirical Study. J. MANAGE. INFO. SYST. 1 (2), 4–17.

Hopenhayn, H. A., 1992. Entry, Exit, and Firm Dynamics in Long Run Equilibrium. Econometrica 60 (5), 1127–50.

Hsieh, C.-T., Klenow, P. J., 2009. Misallocation and Manufacturing TFP in China and India. The Quarterly Journal of Economics 124 (4), 1403–1448.

Idson, T. L., 1989. Establishment Size Differentials in Internal Mobility. The Review of Economics and Statistics 71 (4), 721–24.

Jovanovic, B., 1979. Job Matching and the Theory of Turnover. Journal of Political Economy 87 (5), 972–90.

Kambourov, G., Manovskii, I., 2008. Rising Occupational and Industry Mobility in the United States: 1968-97. International Economic Review 49 (1), 41–79.

Kambourov, G., Manovskii, I., 2009. Occupational Mobility and Wage Inequality. Review of Economic Studies 76 (2), 731–759.

Kircher, P., Manovski, I., Groes, F. N., forthcoming. The U-Shapes of Occupational Mobility. Review of Economic Studies.

Kramarz, F., Postel-Vinay, F., Robin, J.-M., 2014. Occupational Mobility and Wage Dynamics Within and Between Firms. Mimeo.

Lee, G., Xia, W., 2006. Organizational Size and IT Innovation Adoption: A Meta-Analysis . Information & Management 43 (8), 975 – 985.

Lise, J., Postel-Vinay, F., 2014. Multidimensional Skills, Sorting, and Human Capital Accumulation. Mimeo.

Melitz, M. J., Trefler, D., 2012. Gains from Trade When Firms Matter. Journal of Economic Perspectives 26 (2), 91–118.

Moscarini, G., Thomsson, K., 2007. Occupational and Job Mobility in the US. Scandinavian Journal of Economics 109 (4), 807–836.

Papageorgiou, T., 2010. Large Firms and Internal Labor Markets. 2010 Meeting Papers 1216, Society for Economic Dynamics.

Ransom, M. R., Oaxaca, R. L., 2010. New Market Power Models and Sex Differences in Pay. Journal of Labor Economics 28 (2), 267–289.

Rubinstein, A., 1982. Perfect Equilibrium in a Bargaining Model. Econometrica 50 (1), 97–109.

Stijepic, D., 2015a. Globalization, Worker Mobility and Wage Inequality. Available at SSRN: http://ssrn.com/abstract=2635540.

Stijepic, D., 2015b. Job Mobility and Sorting: Theory and Evidence. Available at SSRN: http://ssrn.com/abstract=2635518.

Stijepic, D., 2015c. Workplace Heterogeneity and the Returns to Versatility. Available at SSRN: http://ssrn.com/abstract=2635542.

Tybout, J. R., 2008. Plant- and Firm-Level Evidence on New Trade Theories. In: Handbook of International Trade. Blackwell Publishing Ltd, pp. 388–415.

U.S. Department of Commerce, B. o. t. C., 1950. 1947 Census of Manufactures. Vol. 1, General Summary. Washington, DC: U.S. Government Printing Office.

U.S. Department of Commerce, B. o. t. C., 1957. 1954 Census of Manufactures. Vol. 1, Summary Statistics. Washington, DC: U.S. Government Printing Office.

U.S. Department of Commerce, B. o. t. C., 1971. 1967 Census of Manufactures. Vol. 1, Summary and Subject Statistics. Washington, DC: U.S. Government Printing Office.

U.S. Department of Commerce, B. o. t. C., 1976. 1972 Census of Manufactures. Vol. 1, Subject and Special Statistics. Washington, DC: U.S. Government Printing Office.

U.S. Department of Commerce, B. o. t. C., 1981. 1977 Census of Manufactures. Vol. 1, Subject Statistics. Washington, DC: U.S. Government Printing Office.

U.S. Department of Commerce, B. o. t. C., 1985. 1982 Census of Manufactures. Subject Series. General Summary. Part 2. Industry Statistics by Employment Size of Establishment. Washington, DC: U.S. Government Printing Office.

Appendix A. Data and Auxiliary Statistics - For Online Publication

In this section I describe the data sets used for the figures in the main text and I also present some further empirical results.

Appendix A.1. Economic Census

The Economic Census collects information on the United States' economy once every five years, combining both administrative records and establishment surveys. The scope of the Economic Census has evolved over the years. Since 1992 the industries covered by the program account for 98 percent of the gross domestic product. Earlier censuses included less industries and covered less of the U.S. economy, e.g., 76 percent of the gross domestic product in 1987. Instructions on how to obtain data from the Economic Census are available on the homepage of the United States Bureau of the Census:

`http://www.census.gov/econ/census07`

While the manufacturing sector is traditionally well covered, the coverage of other industries varies substantially over time. Therefore, I rely in particular on the Census of Manufactures. I construct the time series for value added per employee and wages by establishment size class in U.S. manufacturing from tabulations in various Census of Manufactures reports: U.S. Department of Commerce (1950, Chapter 3, p. 97) for the year 1947, U.S. Department of Commerce (1957, Chapter 3, p. 1) for the year 1954, U.S. Department of Commerce (1971, Chapter 2, p. 4) for the years 1958 and 1963, U.S. Department of Commerce (1971, Chapter 2, p. 6) for the year 1967, U.S. Department of Commerce (1976, Chapter 2, p. 68) for the year 1972, U.S. Department of Commerce (1981, Chapter 1, p. 59) for the year 1977, and U.S. Department of Commerce (1985, p. 3) for the year 1982. The series for the years 1987 onwards are directly available in machine readable formats on the website of the Bureau of the Census: "MC87I4-1" for the year 1987, "MC92SF4" for the year 1992, "E9731G4" for the year 1997, "ECN_2002_US_31SG105" for the year 2002, and "ECN_2007_US_31SG3" for the year 2007. Value added in 1947 is omitted from the figures in the main text, since it is only reported in unadjusted terms. Adjusted value added also takes into account value added by merchandising operations, and the net change in finished goods and work-in-process inventories between the beginning and the end of the year.

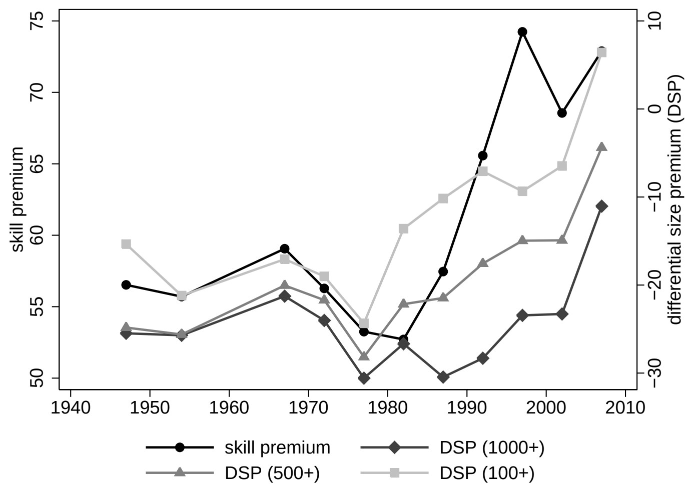

Figure A.5: Skill premium and differential firm size wage premium between skill groups in U.S. manufacturing. The size premium is defined as the wage premium enjoyed by workers at establishments with at least 100, 500 or 1000 employees relative to workers at establishment with less than 100, 500 or 1000 employees, respectively. Non-production workers and production workers serve as proxies for high- and low-skill workers, respectively. Author's calculations based on tabulations from various Census of Manufactures reports.

Figures A.5 depicts the relationship between the skill premium and the differential establishment size wage premium between non-production and production workers for different definitions of large establishments.

Appendix A.2. Current Population Survey

The Current Population Survey (CPS) is administered by the United States Bureau of the Census. Currently, a nationally representative sample of about 65 thousand households are interviewed monthly. Each household is interviewed once a month for four consecutive months, and again for the corresponding four months period a year later, resulting in eight total months in the survey. Each month, new households are added and old ones who complete eight months in the survey are dropped.

Mobility statistics are based on Current Population Survey Basic Monthly data as provided by the National Bureau of Economic Research:

`http://www.nber.org/data/cps_basic.html`

Estimates from May to August 1995 are missing due to changes in the household identification methodology of the Bureau of the Census. CPS sampling weights are used in all calculations.

To ease respondent burden, dependent interviewing - using information from the previous month's interview in the current interview - has been incorporated into the industry and occupation questions from January 1994 onwards. After industry and occupation data are collected in the first month, rather than being asked for the same information every month, individuals interviewed in successive months are asked the following three questions:

1. LAST MONTH, IT WAS REPORTED THAT YOU WORKED FOR (EMPLOYER'S NAME). DO YOU STILL WORK FOR (EMPLOYER'S NAME) (AT YOUR MAIN JOB)?
2. HAVE THE USUAL ACTIVITIES AND DUTIES OF YOUR JOB CHANGED SINCE LAST MONTH?
3. LAST MONTH YOU WERE REPORTED AS (A/AN) (OCCUPATION) AND YOUR USUAL ACTIVITIES WERE (DESCRIPTION). IS THIS AN ACCURATE DESCRIPTION OF YOUR CURRENT JOB?

Only if respondents report changes according to at least one of the three questions, they are asked to specify their occupation and industry status. Otherwise, answers from the previous month are transfered. I exploit the dependent interviewing technique to identify employer–employer transitions and activity changes. In particular, I use the first question to identify employer–employer transitions.[26] I consider the activity to have changed if respondents report "YES" to the second question or "NO" to the third question. For more information on the dependent coding technique used in the monthly CPS survey see Polivka and Rothgeb (1993).

I match the monthly surveys based on the following procedure. First, I generate a preliminary identifier based on the household identifier ("hrhhid"), census state code ("gestcen"), household number ("huhhnum"), gender ("sex"), race (white only, balck only and other using the variables "perace", "prdtrace", and "ptdtrace"), and individual line number ("pulineno") and discard all observations that are not uniquely identified in the monthly cross section. In a second step, I verify longitudinal consistency. I drop all observations if either *(i)* the longitudinal match exhibits an age inconsistency, i.e., the age reported in the following month is not equal to the current age with an error band of -1/+2 years, or *(ii)* the month of interview ("hrmis") exhibits an inconsistency. See Madrian and Lefgren (1999) for further information on longitudinal matching of CPS respondents. Following Fallick and Fleischman (2004), I also exclude all respondents in their first and fifth survey month.

[26] For a detailed discussion on identifying employer–employer transitions using CPS data see Fallick and Fleischman (2004) and Moscarini and Thomsson (2007).

The occupation classification system employed by the Bureau of the Census is changing over time. In particular, in January 2003, the CPS adopted the 2002 Census occupational classification systems; it replaced the 1990 Census occupational classification. I harmonize the classification on the basis of the 1990 occupation classification system following the IPUMS-CPS.

Production and non-production workers in the Census of Manufactures reports are defined as follows:

Production Workers - The number of production workers includes workers (up through the line supervisor level) engaged in fabricating, processing, assembling, inspecting, receiving, storing, handling, packing, warehousing, shipping (but not delivering), maintenance, repair, janitorial and guard services, product development, auxiliary production for plants own use (e.g., power plant), recordkeeping, and other services closely associated with these production operations at the establishment covered by the report. Employees above the working-supervisor level are excluded from this item.

All Other Employees - The other employees number covers nonproduction employees of the manufacturing establishment including those engaged in factory supervision above the line-supervisor level. It includes sales (including driver-salespersons), sales delivery (highway truck drivers and their helpers), advertising, credit, collection, installation and servicing of own products, clerical and routine office functions, executive, purchasing, financing, legal, personnel (including cafeteria, medical, etc.), professional, and technical employees. Also included are employees on the payroll of the manufacturing establishment engaged in the construction of major additions or alterations utilized as a separate workforce.

On the basis of the 1990 Census Classification System I define production and non-production workers in light of the latter definitions as follows. Production workers are individuals in occupations as precision production, craft, and repair occupations (503-699), machine operators, assemblers, and inspectors (703-889), protective service occupations (413-427), and cleaning and building service occupations, except household (448-455). Non-production workers are individuals in occupations as managerial and professional specialty occupations (003-199), technical, sales, and administrative support occupations (203-389), and service occupation, except protective, household, cleaning, and building service occupations (433-447, 456-469). Farming, forestry, and fishing occupations (473-499), private household occupations (403-407), and military occupations (903-905) are excluded.

Appendix B. Derivation of Model Equations - For Online Publication

Appendix B.1. Aggregate Distributions

In this section I use the Fokker–Planck formalism to derive the law of motion for the aggregate distribution of workers over states that is consistent with individuals' laws of motion. A detailed exposition of the Fokker–Planck formalism is provided by Bayer and Wälde (2011). Let $H_i(z,\tau)$, $i \in \{L, H\}$, denote the cumulative distribution of workers over states $z = (p, q, e)$ at time τ. The distribution has various mass points, which I model explicitly, i.e., I assume following decomposition for the density/mass function of the distribution

$$h_i(z,\tau) \equiv \begin{cases} h_i(p,q,e_2,\tau) & \text{if } e = e_2 \\ h_i(p,b,e_1,\tau) & \text{if } e = e_1 \\ h_i(b,b,u,\tau) & \text{if } e = u \end{cases}, \qquad p_{max} \geq p \geq q \geq p_{min}, \tag{B.1}$$

where u indicates unemployed workers, e_1 indicates employed workers who are the first time employed after unemployment and have not yet been able to trigger wage renegotiations, and e_2 indicates employed workers who have triggered wage negotiations at least once while being employed. Furthermore, p indicates the productivity of the firm the worker is employed at, and q indicates the productivity of the firm that triggered the most recent wage renegotiation.

A density to be consistent with the underlying individuals' laws of motions has to satisfy

$$\int \mathbb{E}\left(\frac{\partial \varphi(z(\tau))}{\partial \tau}\right) h_i(z,\tau) dz = \int \varphi(z) \frac{\partial h_i(z,\tau)}{\partial \tau} dz \tag{B.2}$$

for arbitrary functions φ, which are assumed to be differentiable and to have bounded non-zero support, and where $\mathbb{E}(\cdot)$ represents expectations with respect to time. Intuitively, the left-hand side (LHS) describes how the state evolves for individuals in a particular state, whereas the right-hand side (RHS) characterizes how the mass of individuals in a particular state changes. With my decomposition of the distribution the right-hand side (RHS) reads

$$\begin{aligned} RHS = & \int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} \varphi(p,q,e_2) \frac{\partial h_i(p,q,e_2,\tau)}{\partial \tau} dq dp \\ & + \int_{p_{min}}^{p_{max}} \varphi(p,b,e_1) \frac{\partial h_i(p,b,e_1,\tau)}{\partial \tau} dp + \varphi(b,b,u) \frac{\partial h_i(b,b,u,\tau)}{\partial \tau}. \end{aligned} \tag{B.3}$$

Applying a change of variable formula to calculate expectations yields for the left-hand side (LHS) following expression

$$LHS = \zeta_1 + \zeta_2 + \zeta_3, \tag{B.4}$$

where

$$\begin{aligned}\zeta_1 \equiv \int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} \Big\{ \lambda_i \int_{p}^{p_{max}} \varphi(x,p,e_2) - \varphi(p,q,e_2) dF(x) \\ + \lambda_i \int_{q}^{p} \varphi(p,x,e_2) - \varphi(p,q,e_2) dF(x) \\ + \delta_i \left[\varphi(b,b,u) - \varphi(p,q,e_2) \right] \Big\} h_i(p,q,e_2,\tau) dq dp,\end{aligned} \tag{B.5}$$

$$\begin{aligned}\zeta_2 \equiv \int_{p_{min}}^{p_{max}} \Big\{ \lambda_i \int_{p}^{p_{max}} \varphi(x,p,e_2) - \varphi(p,b,e_1) dF(x) \\ + \lambda_i \int_{p_{min}}^{p} \varphi(p,x,e_2) - \varphi(p,b,e_1) dF(x) \\ + \delta_i \left[\varphi(b,b,u) - \varphi(p,b,e_1) \right] \Big\} h_i(p,b,e_1,\tau) dp,\end{aligned} \tag{B.6}$$

and

$$\zeta_3 \equiv \Big\{ \lambda_i \int_{p_{min}}^{p_{max}} \varphi(x,b,e_1) - \varphi(b,b,u) dF(x) \Big\} h_i(b,b,u,\tau). \tag{B.7}$$

Changing the order of integration and relabeling the variables of integration, the terms on the left-hand side (LHS) can be further simplified to

$$\begin{aligned}\zeta_1 = \int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} -(\lambda_i \bar{F}(q) + \delta_i) h_i(p,q,e_2,\tau) \varphi(p,q,e_2) dq dp \\ + \int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} \Big\{ \lambda_i f(p) \int_{p_{min}}^{q} h_i(q,x,e_2,\tau) dx \\ + \lambda_i f(q) \int_{p_{min}}^{q} h_i(p,x,e_2,\tau) dx \Big\} \varphi(p,q,e_2) dq dp \\ + \int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} \delta_i h_i(p,q,e_2,\tau) \varphi(b,b,u) dq dp,\end{aligned} \tag{B.8}$$

$$\zeta_2 = \int_{p_{min}}^{p_{max}} -(\lambda_i + \delta_i) h_i(p,b,e_1,\tau)\varphi(p,b,e_1)dp + \int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} \Big\{\lambda_i f(p) h_i(q,b,e_1,\tau) + \lambda_i f(q) h_i(p,b,e_1,\tau)\Big\}\varphi(p,q,e_2)dqdp + \int_{p_{min}}^{p_{max}} \delta_i h_i(p,b,e_1,\tau)\varphi(b,b,u)dp, \quad \text{(B.9)}$$

and

$$\zeta_3 = \Big\{\lambda_i \int_{p_{min}}^{p_{max}} \varphi(p,b,e_1) f(p)dp - \lambda_i \varphi(b,b,u)\Big\} h_i(b,b,u,\tau). \quad \text{(B.10)}$$

Finally, collecting terms on the left-hand side (LHS) and right-hand side (RHS) results in

$$\int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} \tilde{\zeta}_1(p,q)\varphi(p,q,e_2)dqdp + \int_{p_{min}}^{p_{max}} \tilde{\zeta}_2(p)\varphi(p,b,e_1)dp + \tilde{\zeta}_3\varphi(b,b,u) = 0, \quad \text{(B.11)}$$

where

$$\tilde{\zeta}_1(p,q) \equiv -(\delta + \lambda_i \bar{F}(q)) h_i(p,q,e_2,\tau) + \lambda_i f(p) \int_{p_{min}}^{q} h_i(q,x,e_2,\tau)dx + \lambda_i f(q) \int_{p_{min}}^{q} h_i(p,x,e_2,\tau)dx + \lambda_i f(q) h_i(p,b,e_1,\tau) + \lambda_i f(p) h_i(q,b,e_1,\tau) - \frac{\partial h_i(p,q,e_2,\tau)}{\partial \tau}, \quad \text{(B.12)}$$

$$\tilde{\zeta}_2(p) \equiv -(\delta_i + \lambda_i) h_i(p,b,e_1,\tau) + \lambda_i f(p) h_i(b,b,u,\tau) - \frac{\partial h_i(p,b,e_1,\tau)}{\partial \tau}, \quad \text{(B.13)}$$

and

$$\tilde{\zeta}_3 \equiv \delta_i \int_{p_{min}}^{p_{max}} \int_{p_{min}}^{p} h_i(p,q,e_2,\tau)dqdp + \delta_i \int_{p_{min}}^{p_{max}} h_i(p,b,e_1,\tau)dp - \lambda_i h_i(b,b,u,\tau) - \frac{\partial h_i(b,b,u,\tau)}{\partial \tau}. \quad \text{(B.14)}$$

The equation is obviously satisfied for $\tilde{\zeta}$-terms that are zero for all admissible pq-combinations. This yields the Fokker–Planck equations that describe the evolution of the aggregate distribution of workers over states:

$$\begin{aligned}\frac{\partial h_i(p,q,e_2,\tau)}{\partial\tau} &= -(\delta+\lambda_i\bar{F}(q))h_i(p,q,e_2,\tau)\\ &\quad + \lambda_i f(p)\int_{p_{min}}^{q} h_i(q,x,e_2,\tau)dx + \lambda_i f(q)\int_{p_{min}}^{q} h_i(p,x,e_2,\tau)dx\\ &\quad + \lambda_i f(q)h_i(p,b,e_1,\tau) + \lambda_i f(p)h_i(q,b,e_1,\tau), \end{aligned}\tag{B.15}$$

$$\frac{\partial h_i(p,b,e_1,\tau)}{\partial\tau} = -(\delta_i+\lambda_i)h_i(p,b,e_1,\tau) + \lambda_i f(p)h_i(b,b,u,\tau), \tag{B.16}$$

and

$$\begin{aligned}\frac{\partial h_i(b,b,u,\tau)}{\partial\tau} &= \delta_i\int_{p_{min}}^{p_{max}}\int_{p_{min}}^{p} h_i(p,q,e_2,\tau)dqdp\\ &\quad + \delta_i\int_{p_{min}}^{p_{max}} h_i(p,b,e_1,\tau)dp - \lambda_i h_i(b,b,u,\tau). \end{aligned}\tag{B.17}$$

I restrict the analysis to the steady state of the economy and abstract from any transitional dynamics. Specifically, I consider the equilibrium that is characterized by $\partial h_i(p,q,e_2,\tau)/\partial\tau$, $\partial h_i(p,b,e_1,\tau)/\partial\tau$, $\partial h_i(b,b,u,\tau)/\partial\tau = 0$. Straight forward algebra shows that the following expression for the density satisfies the Fokker–Planck equations in the steady state equilibrium:

$$h_i(z) = \begin{cases} \frac{2\lambda_i^2\delta_i f(p)f(q)}{\left(\delta_i+\lambda_i\bar{F}(q)\right)^3} & \text{if } e=e_2\\ \frac{\delta_i\lambda_i f(p)}{(\delta_i+\lambda_i)^2} & \text{if } e=e_1\\ \frac{\delta_i}{\delta_i+\lambda_i} & \text{if } e=u \end{cases}. \tag{B.18}$$

The latter equation implies following distribution of workers over productivity classes

$$J_i(p) = \frac{\delta_i F(p)}{\delta_i+\lambda_i\bar{F}(p)}, \tag{B.19}$$

and the following unemployment rate

$$u_i = \frac{\delta_i}{\delta_i+\lambda_i}. \tag{B.20}$$

Appendix B.2. Wages

In this section I derive wage Equation (4). The value function of a type i worker employed at a firm of productivity p and earning a wage w is

$$\rho V_i(w,p) = w + \lambda_i \int_{q_i(w,p)}^{p} V_i(\omega_i(x,p),p) - V_i(w,p) dF(x) \\ + \lambda_i \int_{p}^{p_{max}} V_i(\omega_i(p,x),x) - V_i(w,p) dF(x) + \delta_i (U_i - V_i(w,p)), \quad \text{(B.21)}$$

where U_i denotes an unemployed worker's lifetime utility. Using the bargaining equation (3) to substitute for $V_i(\omega_i(p,x),x)$ and $V_i(\omega_i(x,p),p)$, and integrating by parts yields

$$(\rho + \delta_i)\, V_i(w,p) = w + \delta_i U_i + \lambda_i \beta \int_{p}^{p_{max}} \left(\theta_i \frac{\partial V_i}{\partial w} + \frac{\partial V_i}{\partial p} \right) (\theta_i x, x) \bar{F}(x) dx \\ + \lambda_i (1-\beta) \int_{q_i(w,p)}^{p} \left(\theta_i \frac{\partial V_i}{\partial w} + \frac{\partial V_i}{\partial p} \right) (\theta_i x, x) \bar{F}(x) dx. \quad \text{(B.22)}$$

Evaluating at $w = \theta_i p$ and differentiating with respect to p results in

$$\left(\theta_i \frac{\partial V_i}{\partial w} + \frac{\partial V_i}{\partial p} \right) (\theta_i p, p) = \frac{\theta_i}{\rho + \delta_i + \lambda_i \beta \bar{F}(p)}. \quad \text{(B.23)}$$

Inserting the latter two equations into the bargaining equation (3) yields

$$\omega_i(q,p) = \theta_i \left(q + \beta(p-q) - (1-\beta)^2 \lambda_i \int_{q}^{p} \frac{\bar{F}(x)}{\rho + \delta_i + \lambda_i \beta \bar{F}(x)} dx \right). \quad \text{(B.24)}$$

Finally note, that the value function of a unemployed worker reads

$$\rho U_i = b + \lambda_i \int_{p_{min}}^{p_{max}} V(\omega_i(b,x),x) - U_i dF(x), \quad \text{(B.25)}$$

i.e., an unemployed worker enjoys the same lifetime utility as a worker who is employed at a firm of productivity b and appropriates the entire match surplus, $V_i(b,b)$. Therefore, it follows for the wage negotiated with unemployed workers

$$\omega_i(b,p) = \theta_i \left(b + \beta(p-b) - (1-\beta)^2 \lambda_i \int_{p_{min}}^{p} \frac{\bar{F}(x)}{\rho + \delta_i + \lambda_i \beta \bar{F}(x)} dx \right). \quad \text{(B.26)}$$

Combining the employed workers' wage equation (B.24) and the unemployed workers' wage equation (B.26) yields wage equation (4).

Given the steady state distribution of workers over states (B.18), the average wage of a worker of type $i \in \{L, H\}$ at a firm of productivity p is given by

$$\bar{w}_i(p) = \frac{1}{j_i(p)(1-u_i)} \left[\int_{p_{min}}^{p} \omega_i(q,p) h_i(p,q,e_2) dq + \omega_i(b,p) h_i(p,b,e_1) \right]. \quad \text{(B.27)}$$

Straightforward algebra yields equation (5). Similarly, type-i workers' overall average wage reads

$$\begin{aligned} \bar{w}_i = \theta_i \Bigg(\int_{p_{min}}^{p_{max}} w_i(p) j_i(p) dp = \int_{p_{min}}^{p_{max}} \frac{\delta_i f(p) p (\delta_i + \lambda_i)}{(\delta_i + \lambda_i \bar{F}(p))^2} dp \\ - (1-\beta) \int_{p_{min}}^{p_{max}} \frac{\delta_i (\delta_i + \lambda_i)(\rho + \delta_i + \lambda_i \bar{F}(p)) \bar{F}(p)}{(\rho + \delta_i + \lambda_i \beta \bar{F}(p))(\delta_i + \lambda_i \bar{F}(p))^2} dp \\ - (1-\beta_i)(p_{min} - b) \frac{\delta_i}{\delta_i + \lambda_i} \Bigg). \quad \text{(B.28)} \end{aligned}$$

Appendix C. Solution Strategy and Calibration of Transition Parameters - For Online Publication

Appendix C.1. Solution Strategy

Let $s(p) = \lambda_w \bar{F}(p)$ denote the potential job-to-job transition rate at a firm of productivity p. Using the definitions for λ_w and $F(p)$, the equilibrium mass of production facilities at a type p firm is given by $n(p) = -ds/dp(p) \Big/ \gamma(p)\tilde{N}/M$. This yields together with equation (5) following expression for the entrepreneur's first order condition (9)

$$\sum_{i=L,H} \theta_i \left[\int_{p_{min}}^{p} \frac{\alpha_i \delta_i (1-\beta) M_i/M}{(\delta_i + \beta\alpha_i s(q))(\delta_i + \alpha_i s(q))} dq + \frac{(1-\beta)\alpha_i \delta_i (p_{min} - b) M_i/M}{(\delta_i + \alpha_i N/M)^2} \right]$$
$$= \bar{\eta} \left(- \frac{ds/dp\ (p)}{\gamma(p)\tilde{N}/M} \right)^{\eta-1}. \quad \text{(C.1)}$$

Defining $v(p)$ as the left-hand side of the this equation yields equation (15). Taking the first derivative of the left-hand side with respect to p one obtains equation (16).

Equations (15) and (16) are not a system of ordinary differential equations. It is the expression for the productivity density, $\gamma(\cdot)$, that adds additional complexity to the system. Differentiating the expression for the productivity distribution (11) with respect to p yields

$$\gamma(p) = \left(1 - \Phi\left(\frac{\bar{\eta}(\eta-1)}{\eta} \left((v(a_n p)/\bar{\eta})^{\eta/(\eta-1)} - (v(p)/\bar{\eta})^{\eta/(\eta-1)} \right) \right) \right) \tilde{\gamma}(p)$$
$$+ \Phi\left(\frac{\bar{\eta}(\eta-1)}{\eta} \left((v(p)/\bar{\eta})^{\eta/(\eta-1)} - (v(p/a_n)/\bar{\eta})^{\eta/(\eta-1)} \right) \right) \tilde{\gamma}(p/a_n)/a_n, \quad \text{(C.2)}$$

where I substitute with (10) for the cutoff adoption costs, $f_a^*(p)$, before simplifying the expression using equation (C.1) and the definition of $v(p)$. It follows immediately that in order to determine the value of $\gamma(\cdot)$ at a specific point p, one needs to evaluate $v(\cdot)$ at three distinct values, i.e., p/a_n, p, and $a_n p$.

However, the system can be represented as a higher dimensional system of algebraic and ordinary differential equations. I proceed as follows. First, I introduce auxiliary functions $v_j(\cdot)$ and $s_j(\cdot)$ for $j = 1...k$ in order to obtain a piecewise expression for the functions $v(\cdot)$ and $s(\cdot)$:

$$v_j(\tilde{p}) \equiv v(\tilde{p} a_n^{j-1}) \text{ and } s_j(\tilde{p}) \equiv s(\tilde{p} a_n^{j-1}) \text{ for } \tilde{p} \in [p_{min} ... a_n p_{min}]. \quad \text{(C.3)}$$

Therefore, the piecewise-defined expression for $v(\cdot)$ reads

$$v(p) = \begin{cases} v_1(p) & \text{for } p \in [p_{min}...a_n p_{min}] \\ v_2(p/a_n) & \text{for } p \in [a_n p_{min}...a_n^2 p_{min}] \\ \vdots & \vdots \\ v_k(p/a_n^{k-1}) & \text{for } p \in [a_n^{k-1} p_{min}...a_n^k p_{min}] \end{cases} . \tag{C.4}$$

The expression for $s(\cdot)$ is analogous. Furthermore, let $\gamma_j(\cdot)$ be given by

$$\begin{aligned}\gamma_j(\tilde{p}) = &\left(1 - \Phi\left(\frac{\bar{\eta}(\eta-1)}{\eta}\left(\left(v_{j+1}(\tilde{p})/\bar{\eta}\right)^{\eta/(\eta-1)} - \left(v_j(\tilde{p})/\bar{\eta}\right)^{\eta/(\eta-1)}\right)\right)\right)\tilde{\gamma}(a_n^{j-1}\tilde{p}) \\ &+ \Phi\left(\frac{\bar{\eta}(\eta-1)}{\eta}\left(\left(v_j(\tilde{p})/\bar{\eta}\right)^{\eta/(\eta-1)} - \left(v_{j-1}(\tilde{p})/\bar{\eta}\right)^{\eta/(\eta-1)}\right)\right)\tilde{\gamma}(a_n^{j-2}\tilde{p})/a_n. \end{aligned} \tag{C.5}$$

The Pareto density distribution of managerial skills has an unbounded non-zero support. Henceforth, I assume a truncated Pareto distribution instead, i.e., $\hat{\Gamma}(p) \equiv \tilde{\Gamma}(p)/\tilde{\Gamma}(a_n^{k-1} p_{min})$ for $p \in [0...a_n^{k-1} p_{min}]$ and 1 for $p > a_n^{k-1} p_{min}$, to obtain a density distribution of managerial skills with a bounded non-zero support that approximates the original distribution. Therefore, the original system of two differential equations (15) and (16) can be represented as a system of $2k$ ordinary differential equations defined on $\tilde{p} \in [p_{min}...a_n p_{min}]$:

$$v_j(\tilde{p}) = \bar{\eta}\left(-\frac{ds_j/d\tilde{p}\ (\tilde{p})}{a_n^{j-1}\gamma_j(\tilde{p})\tilde{N}/M}\right)^{\eta-1}, \quad \text{and} \tag{C.6}$$

$$dv_j/d\tilde{p}\ (\tilde{p}) = \sum_{i=L,H} \frac{\theta_i \alpha_i \delta_i (1-\beta) M_i/M}{(\delta_i + \beta\alpha_i s_j(\tilde{p}))(\delta_i + \alpha_i s_j(\tilde{p}))}, \tag{C.7}$$

for $j = 1...k$. The latter system of ordinary differential equations requires beyond the original two boundary conditions, i.e $s_k(a_n p_{min}) = 0$ and $v_1(p_{min}) = \sum_{i=L,H}(1-\beta)\theta_i\alpha_i\delta_i(p_{min} - b)M_i/M \Big/ (\delta_i + \alpha_i N/M)^2$, additional $2k-2$ conditions, which are given by $v_j(p_{min}) = v_{j-1}(a_n p_{min})$ for $j = 2...k$ and $s_j(a_n p_{min}) = s_{j+1}(p_{min})$ for $j = 1...k-1$. Finally, it is imposed that $v_0(\tilde{p}) = \bar{\eta}\left(\frac{f_p \eta}{\bar{\eta}(\eta-1)}\right)^{(\eta-1)/\eta}$ and $v_{k+1}(\tilde{p}) = v_k(\tilde{p})$.

Appendix C.2. Calibration of Transition Parameters

The separation rate at a firm of productivity p is given by $\delta_i + \lambda_i \bar{F}(p)$. Integrating over the distribution of workers over firm productivity classes, $J_i(\cdot)$, yields following expression for the unconditional separation rate:

$$\bar{\lambda}_i + \delta_i \equiv \int_{p_{min}}^{p_{max}} \left[\delta_i + \lambda_i \bar{F}(p)\right] dJ_i(p) = \frac{\delta_i(\delta_i + \lambda_i)}{\lambda_i} ln\left(\frac{\delta_i + \lambda_i}{\delta_i}\right) \tag{C.8}$$

Consider a mass $m(0)$ of initially employed workers and let $m(\tau)$ denote the mass of the workers who are employed at time τ as well. The law of motion for $m(\tau)$ is

$$\frac{dm}{d\tau}(\tau) = -\delta_i m(\tau) + \lambda_i(m(0) - m(\tau)). \tag{C.9}$$

Furthermore, let $m|_{f=f_0}(\tau)$ denote the mass of workers who at time τ are still working for the same employer as at time $\tau = 0$. The law of motion for $m|_{f=f_0}(\tau)$ is given by

$$\frac{dm|_{f=f_0}}{d\tau}(\tau) \approx -(\bar{\lambda}_i + \delta_i) m|_{f=f_0}(\tau), \tag{C.10}$$

where the latter equation is only exact for marginal changes in time. Equations (C.9) and (C.10) are linear differential equations and allow following representations

$$m(\tau)/m(0) = e^{-(\delta_i+\lambda_i)\tau} + \frac{\lambda_i}{\lambda_i + \delta_i}\left(1 - e^{-(\delta_i+\lambda_i)\tau}\right) \text{ and} \tag{C.11}$$

$$m|_{f=f_0}(\tau)/m(0) \approx e^{-(\bar{\lambda}_i+\delta_i)\tau}, \tag{C.12}$$

respectively. Given values for $m(\tau)/m(0)$ and $m|_{f=f_0}(\tau)/m(0)$, solving the preceding two equations yields the respective values for δ_i and λ_i.