Economic Growth through Worker Reallocation: The Role of Knowledge Spillovers*

Eero Mäkynen†

September 12, 2024

I explore the role of knowledge diffusion among producers as a driver of productivity growth. By estimating a reduced form model with Finnish employer-employee data, I find evidence that employing workers from more efficient establishments enhances productivity. To understand the aggregate impact of the finding, I develop an endogenous growth model that incorporates the transmission of knowledge through worker reallocation. The calibrated model reveals that knowledge spillovers can increase aggregate productivity growth by 0.19 percentage points and significantly impact output. Moreover, I establish that this mechanism can exacerbate the adverse impact of firing costs on aggregate outcomes.

Keywords: knowledge diffusion, firm dynamics, worker reallocation, economic growth

JEL Codes: D24, E23, E24, J63, O33, O40

---

*I am indebted to Martin Ellison, William Kerr, Petr Sedláček, Vincent Sterk and Oskari Vähämaa for detailed feedback. I would also like to thank Guido Ascari, Vladimir Asriyan, Vasco Carvalho, Jesús Fernández-Villaverde, Shigeru Fujita, Sampreet Goraya, Axel Gottfries, John Haltiwanger, Heikki Kauppi, Jouko Vilmunen, Ludo Visschers, and seminar and conference participants at the University of Oxford, University of Turku, Helsinki GSE and Nordic Summer Symposium in Macroeconomics 2022 for insightful discussions and suggestions. First version: May 2021. Declarations of interest: none.

†University of Turku, FI-20014 Turun yliopisto, Finland (email: eeanma@utu.fi).

1 Introduction

Empirically we observe considerable idiosyncratic variation in the firms' productivities and in the demand each of them faces[1]. This emphasizes the importance of the reallocation of the input factors and suggests that policies interfering with efficient allocation can have significant aggregate consequences[2]. Recent studies by Poschke (2009) and Mukoyama & Osotimehin (2019) have shown effects are not limited to the levels but can also impact the growth path. Traditionally, when exploring how worker flows aggregate shape outcomes, we assume workers to be a resource without memory simply being allocated from one business to another. However, it is likely that workers also diffuse knowledge across establishments, as emphasized by the recent growth literature and the wide use of non-compete contracts[3].

A growing body of empirical research supports the idea of workers transmitting knowledge between producers. Specifically, producers who hire workers from their more productive counterparts appear to experience productivity gains. However, while previous studies have shown that knowledge diffusion can lead to productivity improvements at the firm level, it is unclear whether these improvements translate into aggregate gains for the economy as a whole. Therefore in this paper, I develop and calibrate a general equilibrium model where workers can diffuse knowledge across establishments and show that knowledge diffusion contributes significantly to aggregate outcomes and the effects of firing cost.

First, to explore the relationship between worker mobility and productivity in a reduced form, I propose an extension to the control function approach typically used in production function estimation.

[1]The large dispersion in firm productivities has been pointed out by, e.g., Syverson (2004). Hottman, Redding & Weinstein (2016) show that over half of the firm-size-variation can be attributed to demand heterogeneity.

[2]For example, Haltiwanger, Scarpetta & Schweiger (2014) find an empirical relationship between a high level of employment protection and a low pace of job reallocation. The connection between employment protection and productivity has been analyzed, for example, by Moscoso Boedo & Mukoyama (2012), Da-Rocha, Restuccia & Tavares (2019), Raurich, Sánchez-Losada & Vilalta-Bufí (2015) and Autor, Kerr & Kugler (2007).

[3]The role of knowledge flows between producers has been emphasized, e.g. by Lucas (2009), Lucas & Moll (2014), and Perla & Tonetti (2014). Shi (2023) points out that about 64% of executives in publicly listed firms have signed non-compete contracts.

By making a marginal change to the assumptions, I can use it to estimate spillovers alongside input elasticities. However, the extended method requires additional information on mobility links across producers to identify the average spillover per hire from a more productive establishment. The advantage of this approach is that it addresses the endogeneity issues that could otherwise compromise the estimation and lines up with the quantitative model specified later.

The empirical findings support hiring as a channel of knowledge diffusion in the administrative data on Finnish manufacturing establishments. On average, hiring a worker from more productive establishments is connected with a 0.56 percent increase in productivity, indicating positive spillover effects. The result is robust for alternative specifications and consistent with previous research with other countries, suggesting that Finland's observed connection is not unique.

Next, to explore the aggregate significance of knowledge diffusion through hiring, I develop a model of establishment dynamics by extending an endogenous growth version of Hopenhayn & Rogerson's (1993) model in Poschke (2009) with a knowledge diffusion mechanism. A key feature of the model is that workers changing employers may retain productivity-enhancing knowledge. When establishments dismiss workers due to idiosyncratic productivity shocks or exit decisions, they become available for hire, and other establishments may hire them to expand or replace dismissed workers. Some new workers may have previously worked for a more productive employer and can pass on their knowledge to the new establishment, leading to productivity gains. From the establishments' perspective, the potential for attaining new knowledge presents an opportunity to enhance productivity by hiring an additional worker and incurring the associated costs of adjustment. The likelihood of establishments benefiting from the knowledge of newly-hired workers depends, in part, on their relative position in the productivity distribution.

In the model, the aggregate growth depends on but is not solely defined by knowledge diffusion through worker reallocation. The diffusion directly impacts growth by boosting the mean productivity of incumbent establishments. The rest of the productivity improvements stem from the random-growth mechanism, which operates through productivity shocks.

The shocks increase the variance of establishment productivities, indicating that the productivity of some establishments improves while others are forced under the endogenously determined profitability limit. The increase in the variance and the simultaneous left-truncation of the productivity distribution enhance incumbents' mean productivity, leading to aggregate growth. Entrants play a central role in this process as they imitate incumbents' growing average productivity, thereby sustaining economic growth.

I utilize the model's flexibility to isolate the contribution of knowledge diffusion through hiring to aggregate productivity growth. Simultaneously targeting the micro-level spillover estimate and the aggregate growth rate as part of the internal calibration, the model successfully replicates the targeted reduced-form connection between worker flows from more productive units and establishment productivity growth. Additionally, the internal calibration incorporates central moments of establishment dynamics, such as establishment size, job turnover, and entry rate.

To derive the main results, I compare an economy with knowledge diffusion through hiring to a hypothetical one without it. When workers diffuse knowledge, it enhances the growth of low-productive establishments, thereby increasing the overall mean size of establishments by boosting the left tail of the productivity distribution. These changes, combined with other changing dimensions like entry rate and price adjustments, result in a 0.19 percentage point increase in aggregate productivity growth and a 1.8 percent increase in output. Furthermore, according to the compensating variation, consumption in an economy without spillovers would need to be raised by 2 percent to achieve the same level of lifetime utility. The welfare comparison is helpful in that it takes into account the simultaneous changes in output's level and growth.

My findings show that worker-transmitted knowledge significantly impacts a country's growth rate, highlighting the substantial impact labor market policies can have on a country's growth rate. To illustrate this point, I examine the role of firing costs that equal one year's wage. Introducing such firing cost leads to a 2 percent decrease in output and a 0.14 percentage points reduction in the growth rate. The equivalent variation amounts to a 6 percent decline. To provide a basis for comparison, I recalibrate the model without the spillover mechanism and repeat the same exercise. In this case, the output decreases by 1.6 percent while the growth rate drops by 0.04 percentage points. The compensating variation, which summarizes the changes, indicates that a 4 percent increase in consumption would be required to offset the firing costs. The results demonstrate that spillovers amplify the negative effects of firing costs by a factor of 1.5.

Related Literature. Several studies have explored the connection between firing costs and the level and growth of aggregate productivity. The literature originates from Bentolila & Bertola (1990) and Hopenhayn & Rogerson (1993). Hopenhayn & Rogerson (1993) find that firing costs reduce aggregate productivity, and subsequent literature has focused on understanding the relationship between firing costs and the level of productivity using a variety of empirical and structural approaches. By explicitly studying the effect of firing costs on aggregate productivity growth, Poschke (2009) finds that a firing tax decreases aggregate growth if it applies to all producers. Mukoyama & Osotimehin (2019) find a similar negative growth effect for labor adjustment costs in their calibration, where innovations from entrants primarily drive aggregate growth. I contribute to the firing cost discussion by demonstrating that considering knowledge diffusion through hiring amplifies the negative impact of firing costs on aggregate growth.

Previous literature has examined the role of knowledge transfer between producers as a source of economic growth.[4] For instance, Perla & Tonetti (2014) and Lucas & Moll (2014) explore producers' time allocation decisions between producing and searching for new ideas. Furthermore, Alvarez, Buera & Lucas (2008, 2013), Perla, Tonetti & Waugh (2021), and Buera & Oberfield (2020) have investigated knowledge diffusion in the context of trade. In their models, producers trade goods and disseminate knowledge, resulting in additional positive effects of trade beyond standard efficiency gains from reallocation. In this paper, I incorporate a similar type of endogenous flow of new ideas, which is now influenced by the hiring policies of establishments and the distribution of their productivities. Furthermore, I demonstrate how knowledge diffusion amplifies the gains from reallocation in the context of firing costs.

[4] see, e.g., survey article by Buera & Lucas (2018)

The literature stemming from the seminal contribution of Klette & Kortum (2004) analyzes aggregate growth by examining firms' R&D investment decisions.[5] In contrast to these studies, my model assumes that new technology is generated through a random process and does not incorporate producers' R&D decisions. However, the knowledge diffusion mechanism provides an additional explanation for productivity growth arising from producer choices.

Connecting knowledge flows through worker reallocation to aggregate growth has similarities with studies by Sohail (2021), Baslandze (2022), and Engbom (2023) that examine the dynamics of spinouts, which are firms founded by former employees of incumbents, and with Bradley & Gottfries (2022), who explore the relationship between labor market fluidity and aggregate growth through imitation. In this context, worker mobility is crucial in determining aggregate growth. The key difference is that this strand of literature focuses on the sources of firm heterogeneity at the time of entry, which Sterk ⓡ Sedláček ⓡ Pugsley (2021) have shown to form a significant amount of overall firm heterogeneity. My approach complements this research by concentrating on understanding the differences arising after entry partly attributable to worker-transmitted knowledge.

The possibility of learning through hiring provides individual firms with control over their future productivity. Gabler & Poschke (2013) examine a similar mechanism where firms have control over their productivity through investments in experimentation. However, in their paper, the firms draw the experiment's outcome from an exogenous distribution, distinguishing their work from this paper. In this study, the distribution from which incumbents obtain new technologies is an equilibrium object.

My theoretical framework heavily relies on the fact that workers can convey knowledge between firms. Empirical studies by Parrotta & Pozzoli (2012), Stoyanov & Zubanov (2012), and Serafinelli (2019) have documented the connection between hiring and firms' productivity growth in other countries. Among these, Serafinelli (2019) is most closely related to this paper as it also employs the control function approach introduced in Olley & Pakes (1996), Levinsohn & Petrin (2003), and Ackerberg, Caves &

[5]For recent contributions see e.g. Akcigit & Ates (2021), Acemoglu & Akcigit (2012) and Akcigit & Kerr (2018).

Frazer (2015, hereafter ACF) to address simultaneity and selection issues. An important difference is that, instead of predetermining "good" firms and studying how mobility from them contributes to firm productivity, I extend the method of ACF in such a way that input elasticities and spillovers can be determined simultaneously with only information on mobility links.

As this paper focuses on how knowledge transmits from one establishment to another, with workers serving as mere intermediaries, I have dedicated more details to modeling the establishments' hiring, separation, exit, and entry choices. Consequently, I have abstracted from some aspects of human capital heterogeneity across workers. In studies of Gregory (2020), Jarosch, Oberfield & Rossi-Hansberg (2021), Engbom (2021), and Shi (2023), the relationship between individuals' human capital development and employers' characteristics is examined more thoroughly. However, generally in these type of studies, the firm's productivity component, which enhances the efficiency of all workers in a multi-worker firm context, only evolves exogenously.

2 Empirical Motivation for the Key Mechanism

This section presents empirical evidence on knowledge diffusion through hiring using a matched employer-employee dataset. First, I provide descriptive evidence of the knowledge diffusion across establishments. Next, to estimate spillover effects consistent with the quantitative framework, I propose a method based on the control function approach. The results indicate that hiring from a more productive establishment increases productivity by 0.56 percent on average.

2.1 Evidence on Productivity Spillovers

The purpose of this section is to demonstrate that hiring workers from more productive units can enhance the productivity of an establishment. For the following analysis, I use employer-employee data from manufacturing establishments in Finland between 1995 and 2012. The sample size is around 116 thousand establishment observations from 15 thousand unique establishments. In Appendix A, I give a more detailed description of the data and a thorough explanation of the data-cleaning process.

To be more specific about the relationship between hiring and productivity in the paper, I first give a simple formal representation of a technology process that includes spillover effects, which helps conceptualize the section's results. Define h^+ as the number of employees the establishment hires from its more productive rivals. Then, a productivity process that contains a spillover component can be defined, for example, as follows

$$z_t = \beta_z h^+_{t-1} + \Upsilon(z_{t-1}) + \vartheta_t, \tag{1}$$

where z is the establishment's productivity, ϑ is the productivity shock, and Υ is a function of the productivity of the previous period, e.g. $\Upsilon(z_{t-1}) = \rho z_{t-1}$, where ρ is the persistence. The core component of the equation is the multiplier β_z, which defines the relationship between today's productivity and the hiring decisions made yesterday with yesterday's information. The multiplier can be referred to as the spillover effect because it defines the average productivity impact of hiring from a more productive unit. The specification is only one of several options. However, it can be associated with many quantitative frameworks on firm dynamics and, therefore, serves as a useful starting point. The specification makes it clear that with data on productivity and mobility links between establishments, such a relationship could easily be investigated. However, productivity is not observable, which further complicates the analysis.

The easiest way of proceeding towards reasonable productivity measure, and thus the exploration of spillovers, is to estimate a production function and use it to recover a productivity measure. For this purpose, I follow the control function approach developed by Olley & Pakes (1996) and refined by Levinsohn & Petrin (2003) and Ackerberg et al. (2015). Roughly summarized, the central insight from these studies is that unobserved productivity can be reasonably proxied by a variable expenditure such as materials. This is because flexible inputs are assumed to respond without lag to a productivity shock and, therefore, contain information about productivity that is not contained in the more rigid inputs, such as capital or labor, that are assumed to be predetermined in the current period.

Controlling for unobserved productivity using expenditure on flexible input then allows for consistent estimation of the production function parameters. The method is widely used in research on firm dynamics and, more specifically, in studies of the aggregate consequences of markups, returns to scale, and capital composition.[6] I explain the further details of production function estimation in Appendix B.

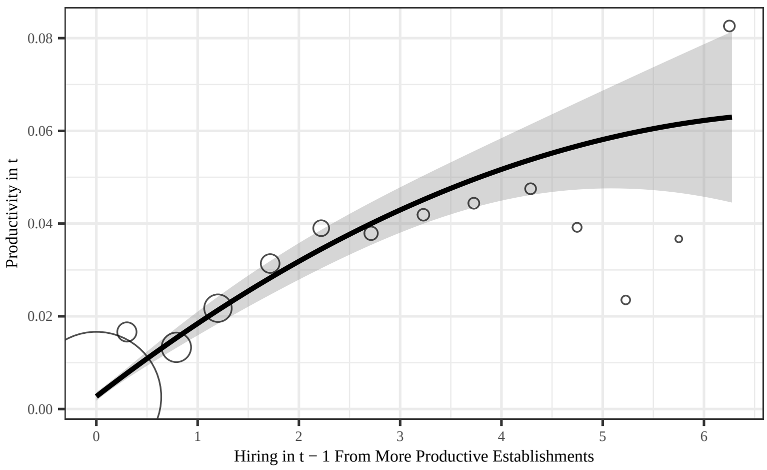
As a by-product of the production function estimation, I retain information on productivity that I use to study spillover effects. Part of the production function estimation is to specify the nature of the producers' productivity process. Typically, the productivity process is assumed to be a first-order Markov process. With a suitable error structure, the productivity process then takes the form

$$z_t = \Upsilon(z_{t-1}) + \omega_t, \tag{2}$$

where ω_t is the productivity shock. In the estimation I need to construct a proxy for z using the flexible input expenditure and, then using the moment conditions generated from the equation above, I can estimate the production function. Thus, I also obtain a value for the productivity shock ω_t and a proxy for productivity through the estimation. By combining the productivity proxy with the mobility links from the employer-employee data, I can calculate how many workers come from more productive units h^+_{t-1} and investigate whether ω_t seems to contain a systematic link to h^+_{t-1}, which I interpret as evidence of spillover effects from new hires.

Using the described approach, I provide the first motivating evidence for hiring as a channel of technology diffusion in Figure 1. The figure shows a clear positive relationship between previous hiring decisions and current productivity for the hiring establishments. We can also see that most establishments did not hire workers from more productive establishments, and those who did hire relatively modest amounts on average. The second-order polynomial also suggests some concavity when the number of hires from better establishments increases. However, the linear trend appears to

[6] Examples of recent papers using ACF estimation are De Loecker, Eeckhout & Unger (2020), Chiavari (2023), and Chiavari & Goraya (2023). The first paper discusses markups, the second focuses on returns to scale, and the final paper examines the structural shift from tangible to intangible capital.



Notes: The figure contains only establishments with positive hiring. The productivity measure is the productivity residual from the ACF estimation. Circles present the number of observations in equispaced intervals on the x-axis.

Figure 1: The relationship between productivity and hiring employees from more productive establishments.

fit inside confidence intervals as well.[7]

When I run a regression between productivity and the previous period's hiring decision for the whole sample, I find further evidence of a statistically significant relationship. Table 1 contains different specifications for the hiring variable. The preferred specification is the one in column 1, where I use the human capital corrected hiring measure. The results show that, on average, hiring from more productive establishments increases productivity by 0.55 percentage points. In column 2, the hiring is measured as raw headcount instead of the human capital corrected one, and, as expected, the multiplier increases (from 0.0055 to 0.0082).

The findings suggest that small and large establishments benefit similarly from hiring workers from more productive establishments. However, two potential mechanisms may limit the gains for larger establishments. First, the likelihood of recruiting employees from highly productive units decreases as the establishment's productivity increases. Second, larger

[7]Note that the number of hires is adjusted based on the average residual wage of new hires from high-productivity establishments, relative to the industry average, in order to account for human capital differences. As a result, hiring an executive is equivalent to hiring an engineer only if their residual wages are the same. Further details on this adjustment are provided in Appendix A, which discusses the data.

establishments may face greater challenges in transmitting new ideas, requiring a larger number of hires from more productive establishments to achieve comparable knowledge transfer. I conduct three additional estimations to see whether the data supports such mechanisms, indicating that larger establishments benefit less and that the largest gains appear to stem from mobility towards smaller and low-productivity establishments.

First, I relate the number of hires to the size of the labor force from the previous period in the specification of column 3. The results indicate that a higher proportion of hires from more productive establishments is associated with productivity gains. Specifically, the multiplier suggests that hiring, e.g., 10 percent of workers from more productive establishments, is linked to an average 0.4 percentage point increase in productivity. Now, as the share of hires from more productive establishments relative to establishment size is generally lower for the larger establishments, the result points towards the spillovers being a more prominent factor for smaller establishments.

Next, I estimate the spillover effect conditional on an establishment's productivity in the previous period, distinguishing between those above and below the mean in column 4. The results show that hiring workers from more productive rivals has a stronger effect on low-productivity establishments. This outcome aligns with the presumption, as lower-productivity establishments appear more likely to benefit from hiring workers from more productive units.

Finally, I assess the spillover effect conditional on the average productivity difference between the hiring establishment and the workers' previous employers. The findings indicate that the larger the average productivity gap between the establishments, the greater the average spillover effect. This suggests that being further from the average productivity level increases the likelihood that random hires will result in spillovers or that targeted hiring from significantly more productive units can be particularly beneficial. In sum, the additional analyses suggest that the relative productivity position of the establishment plays a critical role in moderating the spillover effect.

In the appendix, I also study a spillover measure that is based on Stoyanov & Zubanov (2012). The purpose is to establish a direct connecting

Table 1: Correlation between productivity and hiring from more productive rivals.

	Dependent variable:				
	z_t				
	(1)	(2)	(3)	(4)	(5)
h^+_{t-1}	0.0055	0.0082			
	(0.0021)	(0.0025)			
$\frac{h^+_{t-1}}{n_{t-1}}$			0.0439		
			(0.0104)		
$h^+_{t-1}\mathbb{I}(z_{t-1} \leq \bar{z})$				0.0134	
				(0.0020)	
$h^+_{t-1}\mathbb{I}(z_{t-1} > \bar{z})$				0.0033	
				(0.0017)	
$h^{+\leq 10\%}_{t-1}$					0.0031
					(0.0018)
$h^{+>10\%}_{t-1}$					0.0117
					(0.0019)
3rd ord. polyn. on z_{t-1}	Yes	Yes	Yes	Yes	Yes
human capital adj.	Yes	No	Yes	Yes	Yes

Notes: Robust standard errors in parenthesis. The regression specification follows equation (1) with productivity proxies from the standard ACF estimation. The number of observations is 100,411. The $\bar{z}$ refers to mean of the productivity proxy and percentages in the exponent next to plus sign refer to the productivity difference between sending and receiving establishment in previous year.

point with the existing empirical literature on productivity spillovers. The details of this exercise can be found in Appendix C, where I describe the methodology and the results. To summarize, the results indicate that the spillover effects estimates are the same order of magnitude but a bit lower than those of Stoyanov & Zubanov's (2012) study. The findings give confidence that the results from the Finnish data are within reasonable range in relation to the earlier literature.

So far, I have presented evidence of a correlation between hiring decisions made in the previous period and productivity, using a productivity proxy based on a commonly used empirical approach. The results indicate a positive connection between hiring from more productive rivals, which is more pronounced for small and low-productivity establishments, and that the effect aligns with previous literature. However, there is an inconsistency in the analysis that needs to be addressed. Estimating the productivity proxy requires specifying the productivity process, which affects the estimates themselves. For example, in De Loecker & Warzynski's (2012) paper, productivity is assumed to depend on the productivity of the previous period by a third-order polynomial, which I also used in my analysis. However, if the productivity process includes a systematic component due to spillovers, it must be taken into account to ensure consistent input elasticities and productivity estimates. In the next section, I will discuss how I address this omitted variable issue.

2.2 Consistent Estimation of Spillovers with Control Function Approach

I argue that the spillover coefficient can be consistently estimated as part of the control function approach. A key component of spillover estimation is the proxy for establishment productivity, since the ranking based on productivity determines whether a hire comes from a 'better' establishment. However, this poses a particular challenge: any productivity measure depends on the input elasticities of the production function, and any estimate of input elasticities relies on assumptions about the evolution of productivities, including the potential gains from spillovers. To overcome this challenge, I utilize the flexibility of the control function approach.

I propose an estimation strategy that identifies input elasticities and spillovers jointly. The success of this method depends on two key features of the control function approach. First, the identification is based on the assumption of a first-order Markov productivity process. This assumption allows me to include an additional spillover component in the productivity process that depends on the productivity of the previous period.[8] Second, the control function provides a productivity ranking of establishments that allows for an accurate calculation of worker flows from more productive establishments during the estimation.

[8]This detail makes the control function approach more suitable for my purposes compared to the panel data approaches presented in Chamberlain (1982), Anderson and Hsiao (1982), Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998, 2000), which are another approach to solve the identification problems in estimating the production function.

The goal is to estimate the following set of equations

$$\ln(y_{it}) = \beta_0 + \beta_n \ln(n_{it}) + \beta_k \ln(k_{it}) + z_{it} + \epsilon_{it} \tag{3}$$

$$z_{it} = \Upsilon(z_{it-1}) + \beta_z h^{+}_{it-1}(z_{it-1}, \mathbf{z}_{t-1}) + \vartheta_{it}, \tag{4}$$

where ϵ_{it} is an error term that does not affect the choice of inputs. The variable z_{it} is the time-varying productivity that can influence the producers' input decisions. For the sake of clarity, I refer to hiring from more productive establishments as dependent on the set of productivities in the economy and, more specifically, on those establishment productivities that are linked to the establishment through hiring.

The key and only distinction from the standard control function approach is the quantity h^{+}_{it-1} and its marginal effect β_z. In the estimation, the quantity can be calculated based on the productivity estimates of the lagged period and information about worker flows across establishments. More formally, it is straightforward to note that any productivity proxy, $\hat{z}_{it-1}(\beta_l, \beta_k, \beta_0)$, can be used to construct the set of productivities for recent employers, $\hat{\mathbf{z}}_{t-1}(\beta_l, \beta_k, \beta_0)$, and consequently, $\hat{h}^{+}_{it-1}(\hat{z}_{it-1}(\beta_l, \beta_k, \beta_0), \hat{\mathbf{z}}_{t-1}(\beta_l, \beta_k, \beta_0))$ as the employment today is assumed to be decided on the previous period. Additionally, I need to assume that the establishment's productivity vector, which determines the quality of the employees they hire, is included in the information set. This assumption is consistent with the later-specified model and is no more restrictive than the information assumptions made in models where distributions are treated as state variables. Furthermore, it is necessary to expand the set of instruments to include a suitable variable from the information set of the establishments. Further details of the estimation, I report in Appendix B.

The method described above aims to utilize the benefits of the control function approach for spillover estimation and address potential endogeneity concerns. As I mentioned earlier, the estimation of the production function, which is used to understand productivity evolution, requires including the spillover component to ensure consistency. It's important to note that the ACF method's traditional assumptions eliminate the significant endogeneity concern that arises from mistaking hiring in response to a positive productivity shock as evidence of spillovers.

The estimation assumptions explicitly state that hiring decisions are made before the current period, with full knowledge of how the hiring will affect future productivity, while still being an optimal response to non-spillover-related productivity changes. This is the primary advantage of the approach over other alternatives presented in the literature.

The results from the augmented control function approach in Table 2 reveal that the estimate for the spillovers is positive and statistically significant. The multiplier directly implies that one hire from a more productive establishment brings a 0.56 percent increase in productivity on average. However, the inability to account for price and distributional changes limits our understanding of the aggregate implications of the results. Nonetheless, utilizing the available data, I can perform a back-of-the-envelope calculation by multiplying the spillover estimate with an approximate mean number of hires from the more productive establishments. This calculation yields an average productivity impact of 0.41 percentage points.[9]

For comparison, I have included the estimated production function parameters from the standard ACF approach to Table 2. These were used in the previous section to provide motivating evidence. The labor input elasticity estimates are quite similar, ranging from 0.700 to 0.702. The same applies to the capital elasticities, which range from 0.205 to 0.206. It is important to note that the multipliers display variation. This demonstrates the omitted variable bias that arises if one does not consider the interplay between productivity estimates and spillovers that affect the productivity process. Moreover, the spillover estimate increases compared to the naive approach in the previous section. In Appendix C, I report the consistent estimation results for the alternative specifications, including the specifications from columns 2-5 of Table 1. The results look similar in terms of input elasticities and spillover estimates compared to standard AFC estimation and naive spillover estimation.

[9] As I cannot observe the sending establishment for all hires, I calculate the back-of-envelope figure by dividing the average number of hires from more productive establishments by the share of observable senders, then multiplying it by the regression multiplier. This approach has several limitations as it assumes that all hires are job-to-job transitions from one establishment to another. However, it provides a rough estimate of the potential aggregate impact.

Table 2: Input elasticity estimates and the spillover estimate with extended control function approach.

	Estimation Method:	
	ACF	ACF-SO
n_t	0.702	0.700
	[0.684;0.722]	[0.680;0.720]
k_t	0.206	0.205
	[0.190;0.220]	[0.186;0.218]
h^{+}_{t-1}		0.0056
		[0.0029; 0.0118]

Notes: The numbers in the parenthesis represent the bootstrapped confidence intervals at 5%-level. The bootstrapping is executed by randomizing with replacement the same quantity of establishments, as in the original data, in each round and then repeating the estimation 100 times. In the randomization process, I keep the mobility links intact and recalculate the productivity estimate for the 'sending' establishments that do not end up in the sample based on the current parameter values. The goal is to avoid potential problems arising from the fact that some well-linked establishment does not end up in the bootstrapped sample. The period covered in the analysis is from 1995 to 2012. The instruments used in the estimation include n_{t-1} , n_t, k_{t-1} and k_t. I employ a third-order polynomial to model productivity, in addition to the spillover component, to prevent potential overestimation of spillover effects.

It's worth noting that the spillover estimate doesn't take into account whether the productivity increase comes from individual human capital or the establishment's productivity, which is common to all workers and improves as new knowledge becomes available. Although the human capital adjustment attempts to deal with this, it's important to recognize that the answer isn't definitive. Therefore, I conduct an additional investigation that helps inform the modeling decisions. Specifically, I analyze the impact of a "knowledgeable" worker leaving immediately, who may have shared new information with the establishment. I compare establishments where immediate separation occurs with those establishments where it doesn't, and the results presented in Appendix D indicate no statistically significant difference in evolution of the productivity between the two groups. Note that if there were significant differences, it would indicate that the immediate loss of the knowledgeable worker's transferable human capital would be an important factor for the productivity of an establishment. Thus, it's reasonable to assume, for modeling purposes, that productivity improvements resulting from spillovers increase the productivity of all workers rather than just being embedded in individual workers' transferable human capital. This will greatly increase the tractability of the model and allow dedicating more details to the establishment dynamics.

As worker flows from more productive establishments appear to contribute to the productivity of an establishment, the next step is to assess the quantitative significance of this mechanism. In the following section, I develop a general equilibrium model to achieve this objective. The quantitative framework respects all the timing assumptions made in the above estimation, and I will later use these empirical results to calibrate the model.

3 Model with Knowledge Diffusion through Hiring

To analyze the aggregate significance of knowledge diffusion through hiring, I develop a general equilibrium model that incorporates the diffusion of knowledge through hiring. The model is based on the endogenous growth version of Hopenhayn & Rogerson's (1993) model in Poschke (2009). In addition, I introduce a knowledge diffusion component inspired by the work of Lucas (2009), Perla & Tonetti (2014), and Lucas & Moll (2014). In the model, workers learn about their employer's productivity and can share some of that knowledge when they move to new jobs. The ability of workers to transmit knowledge has important implications for establishments' hiring strategies, as they understand the potential for acquiring new knowledge. These decisions, in turn, have broader implications for aggregate outcomes through the general equilibrium.

3.1 Establishments

Incumbents. Incumbents maximize their expected sum of profits by discounting the future at a rate of $1/(1 + r_t)$. They decide on the number of new workers to hire or lay off and whether to continue. The relevant state variables for the incumbent's decisions are productivity, denoted as $z_t \in (0, \bar{z}_t]$, and the number of employees at the beginning of the period, denoted as n_t. Incumbents use decreasing returns to scale technology with labor as the only input, $f(z_t, n_t) = \exp(z_t)n_t^{\alpha}$, where $0 < \alpha < 1$. I exclude capital from consideration as the focus is solely on the employment aspects of the economy.

In each period, establishments pay a fixed operating cost, $f_{f,t}$, and the wage compensations, $w_t n_t$. Moreover, the establishments are subject to convex relative adjustment costs, expressed as $d(h_t, s_t, n_t) = (f_{a,t}/2)[(h_t + s_t)/\bar{n}]^2\bar{n}$, where h_t denotes hires and s_t separations. Compared to the standard adjustment cost function in the investment literature, where the adjustment is related to the current stock, this specification introduces a minor difference. Here, the adjustment is relative to the average employment stock between periods, denoted as $\bar{n} = 0.5(n_{t+1} + n_t) = n_t + 0.5(h_t + s_t)$ where $n_{t+1} = h_t - s_t + n_t$. The functional form quarantees that for a large establishments it is more costly to hire or separate e.g. 50% of their workforce in comparison to a smaller establishment. Moreover, the specification allows consistent handling of entrants with zero workers.

When deciding on the optimal number of workers to hire, establishments affect their workforce size and shape their future productivity. This link between hiring and productivity arises from the potential for spillover effects that can enhance efficiency. Unlike traditional firm dynamics models, which assume an exogenous productivity process, the model constitutes the endogenous relationship between hiring decisions and future productivity outcomes.

The probability of spillover depends on two endogenous factors: the number of hires, denoted as h_t, and the distribution of knowledge among the pool of reallocating workers, represented by $F_t(z)$. Together, these factors determine the likelihood of achieving a fixed amount of spillover denoted as η. Workers from more productive establishments than z_t can transmit the fixed spillover. However, the transmission of knowledge is uncertain, and I assume that succeeding in the implementation of even one worker's production knowledge is sufficient for receiving the spillover. Therefore, there is a probability of $1 - F_t(z_t)^{h_t}$ that even one knowledge transmission is successful. To summarize, the endogenous component of the establishments' productivity process can be expressed as:

$$\eta\chi_t, \qquad \text{where} \quad \chi_t \sim \text{Bernoulli}(1 - F_t(z_t)^{h_t}). \tag{5}$$

In this equation, χ_t is a stochastic variable that takes on values of zero or one, indicating whether there is a successful implementation of a new worker's knowledge that leads to the occurrence of spillover. I discuss the spillover component of incumbent productivity in greater detail in Section 3.5.

In addition to the endogenous component, idiosyncratic shocks impact the establishments' productivity. The shocks, denoted as u_t, are drawn from a normal distribution with mean zero and variance σ_u^2. The shocks, spillovers, and current productivity collectively determine the next period's productivity

$$z_{t+1} = z_t + \eta\chi_t + u_t, \qquad u_t \sim \mathrm{N}(0, \sigma_u^2). \tag{6}$$

One noteworthy feature of the productivity process is that it exhibits a random walk without the spillover component. The random walk aspect is central to the mechanism that generates the residual growth, along with endogenous exit and entrant imitation, that cannot be attributed to the knowledge diffusion through hiring.

By integrating all the elements, we can formulate the value function for incumbents as follows:

$$\begin{aligned} V(z_t, n_t) = \max_{h_t, s_t, y_t} \Bigg\{ \pi(z_t, n_t) + \frac{1}{1+r_t} \max\{\mathrm{E}_{t|h_t,z_t}[V(z_{t+1}, n_{t+1})], \quad &(7) \\ - c(0, n_t, n_t)\} \Bigg\} \quad &(8) \end{aligned}$$

$$\begin{aligned} \text{s.t.} \quad \pi(z_t, n_t) &= \exp(z_t) n_t^{\alpha} - w_t n_t - w_t d(h_t, s_t, n_t) - w_t f_{f,t} &(9) \\ n_{t+1} &= n_t + h_t - s_t, &(10) \end{aligned}$$

where y_t represents the choice between continuing and exiting, and the optimization problem is subject to knowledge distribution and prices. The solution to the incumbent's problem consists of three policy functions: $h_t(z_t, n_t)$, $s_t(z_t, n_t)$, and $y_t(z_t, n_t)$. These functions describe the employment choices and continuation decisions made by incumbents. Additionally, to ease the notation, I also use $n_{t+1}(z_t, n_t)$ to denote the next period employment that is consistent with the law of motion for labor and above policy functions.

From the standpoint of an incumbent establishment, the diffusion of knowledge through hiring has several implications for their decision-making. Firstly, when hiring is added as a part of the establishment's expected productivity, the optimal size of the establishment can change. In some cases, hiring additional workers may be beneficial to enhance the likelihood of experiencing spillover effects. This mechanism also means that scaling up can occur gradually, with positive hiring over several periods. However, it is important to note that the presence of technology with decreasing returns to scale limits the establishment's ability to expand its size infinitely. These behavioral changes emerge from the knowledge diffusion mechanism, distinguishing this model from a canonical firm/establishment dynamics model.

It's worth noting that hiring and separation must be handled separately instead of implementing a common employment adjustment policy. This is because spillover effects may make it appealing to rotate workers, even if keeping the establishment's size unchanged is optimal. However, the convex adjustment costs ensure that changing all workers every period is not feasible.

Entrants. The economy features an infinite supply of potential entrants who can imitate the average productivity of incumbents as they assess the profitability of entering the market. The entrants compare the expected value of entering the market to the entry costs, $f_{e,t}$. Therefore, we can express the condition for free entry as follows:

$$w_t f_{e,t} \leq \int V(z_t, 0) G_t(\mathrm{d}z), \qquad G_t \sim \mathrm{N}(a_{e,t}, \sigma_z^2). \tag{11}$$

When entrants decide to enter the market, they draw productivity from the distribution G. The distribution is a normal distribution with a fixed variance σ_z^2 and a mean, $a_{e,t}$, which tracks the incumbents' mean productivity from a distance κ. The tracking of mean productivity by entrants represents the imitation process and serves as a key component of the growth mechanism, as it sustains overall economic growth. Further discussion on this feature is provided in subsection 3.5. Additionally, the initial draws generate some knowledge, which spreads via the knowledge diffusion mechanism.

3.2 Household

The economy's infinitely-lived household aims to maximize the lifetime utility through consumption and labor supply decisions. The lifetime utility consists of periodically separable utility functions $u(c_t, l_t) = \theta \ln(c_t) - l_t$, where θ represents the relative utility parameter. When maximizing lifetime utility, household discounts periodical utilities at a rate of β. Moreover, the maximization problem is subject to a budget constraint given by $q_{t+1} + c_t = (1 + r_t)q_t + w_t l_t$. In the budget constraint, the s is the value of shares owned by households, as they own all the shares of the active and entering establishments. The shares yield a periodic return of $r_t s_t$ and hold a value q_t. The periodic returns are equal to the profits generated by establishments in the equilibrium. By solving the households' maximization problem, I obtain an intra-temporal optimality condition $c_t = w_t\theta$ and standard Euler equation $(c_{t+1}/c_t) = \beta(1 + r_{t+1})$, from which $1 + r_{t+1} = (1 + g)/\beta$ every period. The relationship between rates, growth, and the discount rate is used to discount future profits along the growth path.

3.3 Aggregates and Market Clearing Conditions

Establishment Distribution. It is necessary to solve the distribution of establishments so that I can calculate aggregate variables such as output. The distribution is a measure of establishments over $\mathbf{x}_t = [z_t, n_t]$, and it evolves according to a specific law of motion in each period. The first element that describes the evolution of the distribution is the transition matrix $\Phi_t(\mathbf{x}_{t+1}|\mathbf{x}_t, n_{t+1}(\mathbf{x}_t))$. It contains transition probabilities for the incumbent establishments set by the distribution $F_t(z_t)$ and optimal policies. As a distinction from, for example, Hopenhayn & Rogerson's (1993) model, the optimal employment policy also affects transition probabilities on the productivity dimension. By combining the transition matrix with the entry and exit decisions of establishments, I can specify the law of motion for the establishment distribution, $\mu_t(\mathbf{x}_t)$, as

$$\mu_{t+1}(\mathbf{x}_{t+1}) = \int (1 - y_t(\mathbf{x}_t))\Phi_t(\mathbf{x}_{t+1}|\mathbf{x}_t, n_{t+1}(\mathbf{x}_t))[\mu_t(\mathrm{d}\mathbf{x}_t) + m_t\mathbb{I}(n_t = 0)G_t(\mathrm{d}z_t)], \tag{12}$$

where $\mu_t(\mathbf{x}_t)$ is a measure of establishments in $\mathbf{x}_t$ and m_t is the number of entrants.

By definition, the mean productivity of entrants follows the endogenously determined mean productivity of incumbents, and it can be defined as:

$$a_{i,t} = \int z \left(\int \mu_t(\mathrm{d}\mathbf{x}_t)\right)^{-1} \mu_t(\mathrm{d}\mathbf{x}_t). \tag{13}$$

The incumbents' mean productivity then fixes the mean of the productivity distribution $G_t(z)$ as they are connected through equation $a_{e,t} = a_{i,t} - \kappa$.

Workers' Knowledge Distribution. The core part of the knowledge diffusion mechanism is the knowledge distribution of reallocating workers. For simplicity, each reallocating worker remembers their former employer's productivity level. Consequently, the knowledge distribution $F_t(z)$ is constructed by weighting the establishment distribution according to the number of workers reallocating from each productivity level.[10]

As mentioned previously, knowledge distribution plays a crucial role in shaping the choices of individual establishments through general equilibrium. The feedback link makes the Markov chain that describes the evolution of incumbent establishments' productivities interactive in productivity dimension as changes in distribution impact the behavior of establishments, and their behavior further shapes the distribution.[11]

Labor Market Clearing. Households determine the labor supply and establishments' labor demand; these two must coincide in the equilibrium. To recover the household's labor supply, I impose the condition of asset market clearing, which states that $q_{t+1} = q_t = \int V(z_t, n_t)\mu_t(d\mathbf{x}_t)$ in each period. It implies that the household's supply of labor is $l_t = \theta - \pi_t/w_t$,

[10]The employees paying the fixed, entry, and adjustment costs are assumed to be non-mobile from these roles and thus are not taken into account when applying the weighting to the marginal distribution over productivity.

[11]See, for example, Köning et al. (2016) for a theory of innovation and imitation with interactive Markov chain.

where π is the aggregate profit. The supply must equal the demand set by the establishments. By utilizing the establishment measure $\mu(\mathbf{x})$, establishment demand for labor is

$$\bar{n}_t = \int n_t \mu_t(\mathrm{d}\mathbf{x}_t) + f_{f,t} \int \mu_t(\mathrm{d}\mathbf{x}_t) + \int d(n_{t+1}(\mathbf{x}_t), n_t)\mu(\mathrm{d}\mathbf{x}_t) + f_{e,t} m_t. \quad (14)$$

Correspondingly, we can define aggregate profits, $\bar{\pi}$, which are equal to $r_t d_t$, as

$$\bar{\pi}_t = \int \pi(z_t, n_{t+1}(\mathbf{x}_t), n_t)\mu_t(\mathrm{d}\mathbf{x}_t) - w_t f_{e,t} m_t - w_t \int y(\mathbf{x}_t) d(0, n_t, n_t)\mu(\mathrm{d}\mathbf{x}_t). \quad (15)$$

Now equating the defined demand and supply, we get the labor market clearing $\bar{n}_t = l_t$.

3.4 Balanced-Growth Equilibrium

Before defining the balanced-growth equilibrium, I describe the competitive equilibrium of the economy. The competitive equilibrium, where I have normalized the price of the consumption good to unity, consists of sequences of (1) optimal policies, $\{h_t(z_t, n_t),\, s_t(z_t, n_t),\, y_t(z_t, n_t)\}_{t=0}^{\infty}$, of the incumbent establishment (2) wages $\{w_t\}_{t=0}^{\infty}$, (3) establishment distributions $\{\mu_t(z_t, n_t)\}_{t=0}^{\infty}$, and (4) the masses of entrants $\{m_t\}_{t=0}^{\infty}$. These elements satisfy the following conditions: optimal policies solve the incumbent establishment's problem, wages are such that the free entry condition is met, distribution follows its law of motion, and the labor market clears.

In the subsequent analysis, I will solely focus on the balanced growth equilibrium.[12] The balanced growth equilibrium represents a competitive equilibrium in which aggregate productivity, consumption and output, and wages constantly grow at the rate of g. Additionally, the establishment productivity distribution's shape will remain invariant. However, in logs, it will scale up in steps of g every period.

I stationarize the balanced growth equilibrium by transforming growing variables according to $\hat{b}_t = b_t e^{-gt} = b$ and constant variables according to $\hat{k}_t = k_t = k$. This transformation implies that the establishment

[12]More detailed discussion about this type of equilibrium can be found in Poschke (2009).

productivity process will acquire a negative drift equal to g. The negative drift makes the transformed productivity a relative measure of productivity, and, in each period, the establishment's relative position will deteriorate by the amount of g.

3.5 Discussion of the Model Assumptions and the Mechanism of Aggregate Growth

Spillover Component of Incumbents' Productivity. In the model, the extreme assumption that hiring determines the number of samples drawn from the distribution $F(z)$ is intentionally made to simplify the analysis by reducing the number of free parameters. The low number of parameters allows a more straightforward calibration of the spillovers from the regression coefficient obtained in the empirical section. At first glance, reconciling the regression of the empirical section and the assumed process may seem like a challenge as the functional forms do not exactly match. However, this is not the case.

The empirical specification measures the average effect of hiring a worker from a more productive establishment. This average effect could arise from multiple possible data-generating processes, and given that the true underlying process is unknown, some assumptions must be introduced to anchor the analysis. A natural candidate is a linear specification with constant returns to each hire. However, this leads to a problematic implication: unbounded spillovers from hiring, which is unrealistic.

The empirical analysis already hints at the presence of convexity in the spillover effects, suggesting that spillovers are not directly proportional to the number of workers hired. Therefore, I propose an alternative specification that introduces only one additional free parameter. In this alternative, hiring and workers knowledge distribution affect the probability of obtaining a single, potentially larger spillover, rather than incremental benefits from each worker. This assumption ensures that spillovers represent meaningful disruptions that improve an establishment's practices rather than marginal gains from hiring small numbers of part-time workers, for instance.

Under this assumption, the average spillover effect could be the same as in the linear model, but it rises from a different type of process: some establishments may receive no spillover at all, while others receive a slightly larger contribution, η. Furthermore, the process can be aligned with the empirical model as it is straightforward to simulate data from the model and run similar regression on the simulated data as in the empirical section. By doing so, η can be calibrated by matching the regression coefficients, thus ensuring consistency between the full model and the reduced-form empirical model.

Wage Heterogeneity. The main theoretical contribution of this study lies in enhancing a conventional model of firm dynamics by incorporating a spillover component that the available data can directly inform. This allows for a more accurate evaluation of the effects of the diffusion, as it is not the residual of other growth mechanisms. However, it is important to note that I have overlooked the intricate interplay between wage heterogeneity and knowledge spillovers to improve tractability.

Studying the combination of spillovers and wage heterogeneity is an intriguing avenue of research. For instance, the spillovers could be introduced into an on-the-job search framework based on Postel-Vinay & Robin (2002). Nevertheless, models falling under this category primarily emphasize the employer side of the economy, devoting less attention to some of the central aspects of firm dynamics, such as aggregate growth, multi-worker firms, or endogenous entry and exit. To comprehensively address all these factors simultaneously, I have shifted the model's focus entirely to the producer side to provide an alternative angle.

Growth Mechanism. As the focus is on understanding the aggregate growth consequences of knowledge diffusion, I describe the operation of the growth mechanism in detail. Knowledge diffusion through hiring plays a significant role in determining the rate of economic growth. The mechanism leads to an increase in average productivity as some establishments gain knowledge spillover in each period. However, the probability of the increase depends on the type of worker reallocation in the economy through the equilibrium.

In addition to knowledge diffusion across incumbents, aggregate growth results from establishment selection and idiosyncratic shocks. Since the productivity process does not revert to its mean, the variance of the establishment productivity distribution increases each period. As a result, some establishments are driven to the exit threshold, which truncates the productivity distribution from the left. The truncation and the increase in variance imply an increase in average productivity.

Entrants sustain aggregate growth by imitating the productivity increases generated by the knowledge diffusion mechanism, idiosyncratic shocks, and selection. Without the imitation mechanism, the productivity distribution of establishments would thin out over time, which makes imitation an essential part of the growth mechanism.

4 Quantitative Results

In this section, I assess the quantitative significance of the knowledge diffusion mechanism. I calibrate the model by utilizing the estimates from the empirical section and central moments of establishment dynamics. Based on the calibrated model, knowledge diffusion through hiring increases aggregate productivity growth by 0.19 percentage points. Moreover, knowledge diffusion increases the output by 1.8 percent and consumption equivalent variation combining the growth and level effects is 2 percent. As a policy experiment, I study the interaction of knowledge diffusion and firing costs. Results imply that the adverse effects of firing costs increase by a factor of 1.5.

4.1 Model Calibration

I calibrate the parameters to fit the model to the same data as in the empirical section. The goal is to utilize the empirical section as much as possible in fixing the parameter values. The remaining parameters that cannot directly be attached with external evidence, I calibrate internally to match central moments of the establishment dynamics.

First, I set the discount rate and the utility function parameter. Annual calibrations typically use 0.95 as the discount rate, β, and I follow this convention. The utility function parameter, θ, fixes the aggregate expenditure because the labor supply is fully elastic. I will normalize its value equal to unity, as I am interested in the relative figures of aggregate

Table 3: Model calibration and data fit

Variable	Value	Explanation	Data	Model
External				
β	0.95	Convention		
α	0.70	ACF estimation		
σ_z	0.31	Std. of entrant prod.		
σ_u	0.14	Std. of prod. diff.		
ψ	0.11	Knowledge mob. adj.		
Internal				
η	0.010	Infered from ACF estimate	0.0056	0.0055
f_e	3.6	Infered from mean size	20	20
f_f	0.32	Infered from entry rate	0.067	0.066
f_a	3.2	Infered from job turnover	0.12	0.12
κ	0.32	Infered from growth rate	0.026	0.026
Non-targeted Moments				
		Mean number of "better" hires	0.076	0.060
		Mean size of hires	3.0	1.8
		Std. of hires	9.2	15
		Mean size of layoffs	3.3	1.8
		Std. of layoffs	14	18

variables rather than absolute levels. Similarly to Poschke (2009), I set the upper bound of the grid on z to correspond some large value which indicates establishment size that is extremely rare in the data. In the calibrated model the upper bound implies establishment size of approximately 5000, in the data under 10 establishments.[13]

Second, the estimation of the empirical section gives direct values for the output elasticity of labor, the standard deviation of entrant productivities, and productivity shocks. I use 0.70 from the spillover augmented control function estimation as the output elasticity of labor. Moreover, the control function approach gives an estimate of the productivity level, and I use this information directly in the calibration. To fix the standard deviation of entrants' productivities, I calculate the dispersion of productivities for establishments in ages 0-2, and the corresponding value is 0.31. The law of motion for productivity directly gives a way to back out the variance of productivity shocks net of the spillovers. I use these values and set the

[13]Reporting the exact number is prohibited by Statistics Finland.

standard deviation of productivity innovation to 0.14.

Third, a limitation of the empirical part is that some of the hires come from elsewhere than other manufacturing establishments, and in the model, mobility comes only from other establishments. To avoid biased conclusions, I must apply a correction to this. Therefore, I introduce a parameter ψ, which tells how many reallocating workers remember the previous employers' productivity and, thus, represent the within-manufacturing sector reallocation. I fix the value of this knowledge mobility adjustment parameter based on the ratio of observed manufacturing establishment transitions compared to all hires; the value is 0.11.[14]

Fourth, I use five parameters to target five data moments. The remaining parameters are the spillover size, the entry cost, the fixed cost, the adjustment cost, and the imitation distance parameter. Since the model features decreasing returns from spillovers through hiring, the estimate itself cannot be used directly as it only gives the average spillover per hire. Therefore, I run the same regression for productivity in the simulated data as in the empirical section and match the multipliers β_z using η. The moment matching makes the simulated data indistinguishable from the actual data regarding spillovers, even if the functional forms do not match perfectly. The rest of the calibration is relatively standard; the average size and entry rate are used to find proper values for entry and fixed costs. Job turnover, calculated as the sum of job creation and destruction relative to the total number of jobs, helps determine the adjustment cost parameter. Finally, the average aggregate growth rate between 1995 and 2012 is used to determine the value for the imitation distance parameter. Of course, these arguments are heuristic, as moment matching determines all parameters jointly.

Table 3 displays calibrated parameter values and the empirical fit of the model. From the table, we can observe that entering the manufacturing sector costs the equivalent of a wage paid for 3.6 years to a single worker. Running the establishment entails the cost of a wage paid for approximately 4 months to a single worker. The adjustment cost parameters tell us that rotating a single worker in a mean-sized establishment costs around 4

[14]This is equivalent to saying that $F(z_t)$ has a mass of $1 - \psi$ of zeros that denote workers whose origin cannot be tracked.

months' wage. Overall the empirical fit is good, and the most important characteristics, namely the aggregate growth and the spillover coefficient, give close to an exact match.

To assess the model's ability in capturing non-targeted moments, I have added worker mobility statistics to Table 3. The model successfully replicates the average number of hires from higher-quality establishments. Additionally, the first two moments of hiring and separation align reasonably well with the empirical data. While the model tends to produce slightly lower average hires and separations, it generates a larger standard deviation. This is a reasonable outcome given that the model relies on a single convex adjustment cost function, which symmetrically governs both hiring and separations.

4.2 Quantitative Significance of Knowledge Diffused by Workers

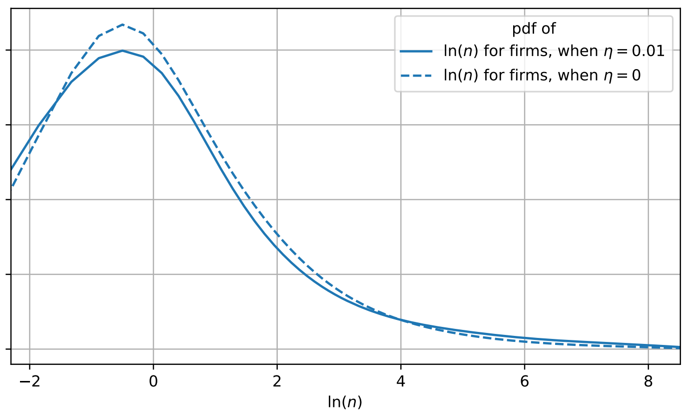
To explore the quantitative significance of knowledge diffusion, I conduct a simple counterfactual experiment by shutting down the spillover channel. To do this, I set η to zero and solve for the balanced growth equilibrium of the model while keeping other parameters fixed. Although this consideration is more theoretical, it mimics a situation where strong enough non-competition clauses could be implemented to shut down any flow of information through workers across establishments.

The main results, reported in Table 4, reveal that knowledge diffusion through hiring has a quantitatively significant impact. Two key observations support the conclusion. First, shutting down the knowledge diffused by workers decreases the aggregate growth rate by 0.19 percentage points. Second, the level of output reduces significantly by 1.8 percent. To summarize the total effect of both changes, I calculate the equivalent variation between the economies. From the household's problem, it is straightforward to determine how much consumption should be increased in the spilloverless economy to achieve the benchmark economy's utility level. The results show that consumption had to be increased by 2 percent to compensate for the lack of productivity increases from the spillovers.

In addition to the growth and welfare effects, the spillover mechanism

Table 4: The aggregate impact of spillovers.

	Calibrated (BM), $\eta = 0.010$	No spillovers, $\eta = 0$
Output / Wage	100	98.2
Growth (%)	2.56	2.37
Labor Demand	100	95.8
Firm Mass	100	142
Entry rate (%)	6.63	6.11
Job Turnover (%)	12.3	12.1
Worker Turnover (%)	15.5	12.1
Mean Size	20.1	12.5
Equivalent Variation	100	102



Notes: The figure contains gaussian kernel density estimates for the marginal distribution over $\ln(n)$. By using the probability mass function implied by $\mu(\mathbf{x})$, I create simulated data of 100000 observations from the marginal distribution and then fit the kernel to this data. As a bandwith, I use 0.25.

Figure 2: Establishment size distributions.

also significantly impacts the central parts of establishment dynamics. Although the effect is small, the spillover channel increases the establishment's job turnover. The more significant change is in worker turnover, as in the model, the spillover mechanism creates incentives to rotate some of the workers, and therefore, job and worker turnover differ under spillovers.[15] The mechanism also increases exit and entry, making the environment more dynamic. However, the mechanism also significantly

[15]Worker turnover is defined as the sum of hirings and separations relative to the total number of jobs.

Table 5: Establishment sizes and worker churning.

	Calibrated (BM), $\eta = 0.010$		No spillovers, $\eta = 0$		
Productivity	n^*	$\frac{h(z,n^*)+s(z,n^*)}{n^*}$	n^*	$\frac{h(z,n^*)+s(z,n^*)}{n^*}$	$\Delta \ln(n)$
0.64	2.5	0.03	3	0	0.17
1.24	25	0.14	20	0	-0.21
1.89	250	0.15	179	0	-0.33
2.68	2500	0.04	2295	0	-0.09

Notes: The productivity levels are attained by using interpolation on policy functions to find out reasonable bencmark establishment sizes. Then, these productivities are re-explored with help of interpolation in the new equilibrium. Also the fact that without spillovers there is no point of churning at the optimal size is used in the last column.

increases the mean size of establishments, despite their shortened expected lifetime. This increase is a natural consequence of the spillover mechanism, as it favors small establishments that can benefit from the knowledge of nearly every worker. The unequal benefits across establishment distribution can also be seen from the whole distribution, which widens in Figure 2. The changes in the size distribution indicate that spillovers let some establishments escape the exit threshold with significant leaps, which would not be possible without the spillover mechanism.[16]

The aggregate results show that the most significant adjustments for establishments occur in worker turnover (churning) and establishment size. Therefore, I present a more detailed analysis of these dimensions in Table 5. The table identifies four representative productivity levels within the benchmark economy, which correspond to reasonable establishment sizes. I then calculate the relative amount of worker-churning for each of these establishments. The analysis reveals that worker turnover is most pronounced in mid-sized establishments. These establishments are sufficiently large to absorb the costs associated with churning while still benefiting from potential positive spillovers, as the distribution of workers' knowledge retains some probability mass on the right tail. In contrast, smaller establishments are deterred by the costs of churning, opting for a more conservative level of worker rotation. Among larger establishments,

[16]In Appendix E, I also report the marginal distribution of productivity in both scenarios and the knowledge distribution of reallocating workers.

the advantages of worker turnover diminish, leading them to similarly adopt lower levels of churning. When the spillover mechanism is removed, the most substantial changes in establishment size are observed in the middle of the distribution. This outcome is expected, as the spillover mechanism disproportionately benefits low to medium-productivity establishments.

The findings indicate that in a relatively undynamic environment like the Finnish manufacturing sector, a significant portion of the growth comes from worker-transmitted knowledge. These results suggest further investigating the impact of worker-based knowledge diffusion. While my analysis has focused on the positive aspects of knowledge diffusion, the following section takes a normative approach to explore how hiring-based knowledge diffusion affects the impact of firing costs. The effect of employment protection legislation, represented by firing costs, has been extensively studied in the literature. Therefore, it's intriguing to see how knowledge diffusion alters the effects of this policy.

4.3 Firing Costs and Knowledge Spillovers

I consider the role of firing costs to see how knowledge diffusion through hiring affects economic policy. Based on the previous literature, the firing costs severely impact output and growth in settings without knowledge diffusion through hiring. In what follows, I show how the impact of firing cost changes when we add the knowledge diffusion mechanism into the mix. Given that the mechanism of interest operates through worker reallocation, any friction limiting workers' movement also impairs knowledge diffusion.

I conduct a counterfactual exercise similar to the previous section to understand how knowledge diffusion changes the effects of firing costs. To the model, I add a firing cost function $d_f(s_t) = \lambda s_t$ also paid in labor and set the value λ to one as in Poschke (2009). Moreover, as the idea is to understand the potential bias in conclusions derived from a model without the spillover mechanism, I re-calibrate the model and exclude the spillover target. Appendix F contains the parameter values for the re-calibrated model. In the re-calibrated model, I repeat the same exercise of increasing the firing costs to one. Comparing the results between these two firing cost increases helps us understand spillovers' role in the consequences of these

Table 6: The aggregate impact of firing costs.

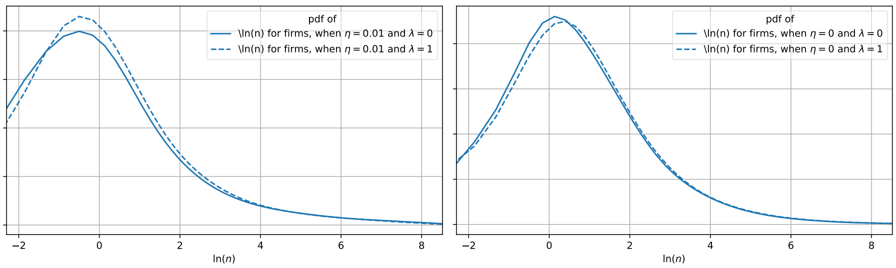
	With spillovers, $\eta = 0.010$		Without spillovers, $\eta = 0$	
	Benchmark, $\lambda = 0$	Firing Costs, $\lambda = 1$	Benchmark, $\lambda = 0$	Firing Costs, $\lambda = 1$
Output / Wage	100	98.0	100	98.4
Growth (%)	2.56	2.42	2.56	2.52
Labor Demand	100	98.0	100	99.4
Firm Mass	100	109	100	91.7
Entry rate (%)	6.63	6.09	6.85	6.63
Job Turnover (%)	12.3	9.71	12.0	8.54
Worker Turnover (%)	15.5	10.3	12.0	8.54
Mean Size	20.1	16.6	20.1	21.1
Equivalent Variation	100	106	100	104

economic policies.

Firing costs impact the central descriptives of the calibrated economy based on the results reported in Table 6. According to the results, the firing costs decrease the aggregate growth by 0.14 percentage points and the output level by 2 percent. Moreover, the equivalent variation between the economies is a 6 percent increase in consumption to compensate for the negative impact of firing costs. Introducing the firing costs also impacts the business dynamism as the turnover decreases by three and a half percentage points, mean size decreases by 3.5 workers, and entry decreases by 0.5 percentage points.

In the hypothetical economy without knowledge spillovers, the firing costs cause smaller changes as in the calibrated economy. According to the results in Table 6, firing costs have 0.04 percentage point effect on aggregate growth. However, the firing cost decreases output by 1.6 percent, and the equivalent variation amounts to 4 percent of consumption. The effect on the business dynamism is similar to before the turnover decreases by five percentage points, the mean size increases by 1 workers, and exit decreases by 0.22 percentage points. By comparing the effect of firing costs between the two alternative specifications, we can see that knowledge diffusion through hiring makes the adverse effects 1.5 times as large.

To examine how establishment heterogeneity responds to separation costs, I plot employment distributions, similar to those presented in



Notes: The figure contains gaussian kernel density estimates for the marginal distribution over $\ln(n)$. By using the probability mass function implied by $\mu(\mathbf{x})$, I create simulated data of 100000 observations from the marginal distribution and then fit the kernel to this data. As a bandwith, I use 0.25.

Figure 3: Distributions of establishment sizes. The left figure represents the model with knowledge diffusion when firing costs are introduced. The figure on the right plots the same excercise in a model that has no spillovers.

the previous section, for the firing cost experiment in Figure 3.[17] The results indicate that firing costs compress the employment distribution in an environment with spillovers by suppressing the mechanism through which establishments can enhance productivity. In contrast, when comparing these outcomes to the standard model without spillovers, we observe a markedly different response. Without spillovers, firing costs increase establishment sizes, as the costs of scaling down hinder downsizing, and wages decline. Overall, these findings demonstrate that predictions regarding establishment sizes, and more broadly, the size distribution of establishments, differ significantly when knowledge spillovers are incorporated into the model.

The size of the growth effect aligns with the values found in the literature. Even if there is no comparable analysis of a similar mechanism, we can compare the effect's size to the previously studied connection between firing costs and growth. Poschke (2009) finds that similar firing costs reduce the growth by 0.09 percentage points, and Mukoyama & Osotimehin (2019) find a 0.1–0.2 percentage point effect of firing costs depending on the calibration. Compared to both studies, the effect I find is within a reasonable range. Notice also that in my model, the adjustment costs already affect the labor adjustments, and the firing cost comes on top of that and introduces the

[17] Appendix E plots the marginal distributions for productivity.

inaction region to the establishments' labor adjustment policies.

5 Conclusion

I find empirical evidence that shows a link between hiring from more productive establishments and the productivity growth of establishments. Motivated by the evidence, I examine the quantitative significance of knowledge diffusion through hiring in a standard firm dynamics framework. I demonstrate that the knowledge diffusion mechanism significantly impacts aggregate growth and establishment dynamics by calibrating the framework to central data moments. Furthermore, I show that the mechanism exacerbates the adverse effects of firing costs.

From a policy perspective, the study shows that firing costs can have a more detrimental impact on aggregate outcomes than previously thought. In addition to hindering reallocation, firing costs reduce the rate of knowledge diffusion. Considering this mechanism reveals that the adverse effect of firing costs can be 1.5 times greater. While the analysis does not explicitly examine the impact of non-compete contracts, it suggests the potential upper limit of their effect if they were to obstruct knowledge flow through worker mobility completely. A more comprehensive analysis of the impact of such contracts on aggregate growth would be an intriguing avenue for future research, which would need to consider the incentives that intellectual property protection creates for innovation.

Throughout my paper, I focus on the establishment dynamics aspect of knowledge diffusion through hiring, providing less detail on modeling the labor market. A more thorough model of the labor market could offer additional insight into the effects of knowledge diffusion on aggregate growth and business dynamics. Additionally, studying the relationship between hiring and producers' productivity growth across a more extensive range of countries could deepen our understanding, as the current evidence is mainly from Nordic countries.

References

Acemoglu, D. and U. Akcigit (2012). Intellectual Property Rights Policy, Competition and Innovation. *Journal of the European Economic Association 10*(1), 1–42.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification Properties of Recent Production Function Estimators. *Econometrica 83*(6), 2411–2451.

Akcigit, U. and S. T. Ates (2021). Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory. *American Economic Journal: Macroeconomics 13*(1), 257–98.

Akcigit, U. and W. R. Kerr (2018). Growth through Heterogeneous Innovations. *Journal of Political Economy 126*(4), 1374–1443.

Alvarez, F. E., F. J. Buera, and R. E. Lucas (2008). Models of Idea Flows. *National Bureau of Economic Research Working Paper Series No. 14135.*

Alvarez, F. E., F. J. Buera, and R. E. Lucas (2013). Idea Flows, Economic Growth, and Trade. *National Bureau of Economic Research Working Paper Series No. 19667.*

Anderson, T. and C. Hsiao (1982). Formulation and Estimation of Dynamic Models using Panel Data. *Journal of Econometrics 18*(1), 47–82.

Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies 58*(2), 277–297.

Arellano, M. and O. Bover (1995). Another Look at the Instrumental Variable Estimation of Error-components Models. *Journal of Econometrics 68*(1), 29–51.

Autor, D. H., W. R. Kerr, and A. D. Kugler (2007). Does Employment Protection Reduce Productivity? Evidence from US States. *The Economic Journal 117*(521), 189–217.

Baslandze, S. (2022). Entrepreneurship through Employee Mobility, Innovation, and Growth. *Federal Reserve Bank of Atlanta Working Paper Series 2022-10*.

Bentolila, S. and G. Bertola (1990). Firing Costs and Labour Demand: How Bad is Eurosclerosis? *The Review of Economic Studies 57*(3), 381–402.

Blundell, R. and S. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics 87*(1), 115–143.

Blundell, R. and S. Bond (2000). GMM Estimation with Persistent Panel Data: An Application to Production Functions. *Econometric Reviews 19*(3), 321–340.

Bradley, J. and A. Gottfries (2022). Labor Market Dynamics and Growth. *Working Paper*.

Buera, F. J. and R. E. Lucas (2018). Idea Flows and Economic Growth. *Annual Review of Economics 10*(1), 315–345.

Buera, F. J. and E. Oberfield (2020). The Global Diffusion of Ideas. *Econometrica 88*(1), 83–114.

Chamberlain, G. (1982). Multivariate Regression Models for Panel Data. *Journal of Econometrics 18*(1), 5–46.

Chiavari, A. (2023). Customer Accumulation, Returns to Scale, and Secular Trends. *Working Paper*.

Chiavari, A. and S. Goraya (2023). The Rise of Intangible Capital and the Macroeconomic Implications. *Working Paper*.

Da-Rocha, J.-M., D. Restuccia, and M. M. Tavares (2019). Firing Costs, Misallocation, and Aggregate Productivity. *Journal of Economic Dynamics and Control 98*, 60–81.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The Rise of Market Power and the Macroeconomic Implications. *The Quarterly Journal of Economics 135*(2), 561–644.

De Loecker, J. and F. Warzynski (2012). Markups and Firm-Level Export Status. *American Economic Review 102*(6), 2437–71.

De Ridder, M., B. Grassi, and G. Morzenti (2022). The Hitchhiker's Guide to Markup Estimation. *CEPR Discussion Paper No. 17532*.

Engbom, N. (2021). Labor Market Fluidity and Human Capital Accumulation. *Working Paper*.

Engbom, N. (2023). Misallocative Growth. *Working Paper*.

Gabler, A. and M. Poschke (2013). Experimentation by Firms, Distortions, and Aggregate Productivity. *Review of Economic Dynamics 16*(1), 26–38.

Gregory, V. (2021). Firms as Learning Environments:Implications for Earnings Dynamics and Job Search. *Working Paper*.

Haltiwanger, J., S. Scarpetta, and H. Schweiger (2014). Cross Country Differences in Job Reallocation: The Role of Industry, Firm Size and Regulations. *Labour Economics 26*, 11–25.

Hopenhayn, H. and R. Rogerson (1993). Job Turnover and Policy Evaluation: A General Equilibrium Analysis. *Journal of Political Economy 101*(5), 915–938.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016). Quantifying the Sources of Firm Heterogeneity. *The Quarterly Journal of Economics 131*(3), 1291–1364.

Jarosch, G., E. Oberfield, and E. Rossi-Hansberg (2021). Learning From Coworkers. *Econometrica 89*(2), 647–676.

Klette, T. and S. Kortum (2004). Innovating Firms and Aggregate Innovation. *Journal of Political Economy 112*(5), 986–1018.

König, M. D., J. Lorenz, and F. Zilibotti (2016). Innovation vs. Imitation and the Evolution of Productivity Distributions. *Theoretical Economics 11*(3), 1053–1102.

Levinsohn, J. and A. Petrin (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies 70*(2), 317–341.

Lochner, B. and B. Schulz (2024). Firm productivity, wages, and sorting. *Journal of Labor Economics 42*(1), 85–119.

Lucas, R. E. (2009). Ideas and Growth. *Economica 76*(301), 1–19.

Lucas, R. E. and B. Moll (2014). Knowledge Growth and the Allocation of Time. *Journal of Political Economy 122*(1), 1–51.

Moscoso Boedo, H. J. and T. Mukoyama (2012). Evaluating the Effects of Entry Regulations and Firing Costs on International Income Differences. *Journal of Economic Growth 17*(2), 143–170.

Mukoyama, T. and S. Osotimehin (2019). Barriers to Reallocation and Economic Growth: The Effects of Firing Costs. *American Economic Journal: Macroeconomics 11*(4), 235–70.

Olley, G. S. and A. Pakes (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica 64*(6), 1263–1297.

Parrotta, P. and D. Pozzoli (2012). The Effect of Learning by Hiring on Productivity. *The RAND Journal of Economics 43*(1), 167–185.

Perla, J. and C. Tonetti (2014). Equilibrium Imitation and Growth. *Journal of Political Economy 122*(1), 52–76.

Perla, J., C. Tonetti, and M. E. Waugh (2021). Equilibrium Technology Diffusion, Trade, and Growth. *American Economic Review 111*(1), 73–128.

Poschke, M. (2009). Employment Protection, Firm Selection, and Growth. *Journal of Monetary Economics 56*(8), 1074–1085.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica 70*(6), 2295–2350.

Raurich, X., F. Sánchez-Losada, and M. Vilalta-Bufí (2015). Knowledge Misallocation and Growth. *Macroeconomic Dynamics 19*(7), 1540–1564.

Serafinelli, M. (2019). "Good" Firms, Worker Flows, and Local Productivity. *Journal of Labor Economics 37*(3), 747–792.

Shi, L. (2023). Optimal Regulation of Noncompete Contracts. *Econometrica 91*(2), 425–463.

Sohail, F. (2021). From Employee to Entrepreneur: Learning, Employer Size, and Spinout Dynamics. *Journal of Economic Dynamics and Control 133*, 104270.

Sterk, V., P. Sedláček, and B. Pugsley (2021). The Nature of Firm Growth. *American Economic Review 111*(2), 547–79.

Stoyanov, A. and N. Zubanov (2012). Productivity Spillovers across Firms through Worker Mobility. *American Economic Journal: Applied Economics 4*(2), 168–98.

Syverson, C. (2004). Market Structure and Productivity: A Concrete Example. *Journal of Political Economy 112*(6), 1181–1222.

Appendix

Appendix A. Finnish Employer-employee Data

In the paper, I use matched employer-employee data from Finnish manufacturing from 1995 to 2012 provided by Statistics Finland. The dataset contains information on all Finnish individuals and their employers, allowing me to track worker movements and identify employer characteristics. The employer data concerns firms in the manufacturing sector with at least 20 employees, including establishments even with fewer workers. I exclude government-owned establishments and special legal forms of companies[18], which leaves me with private sector sole proprietors and limited liability companies. Additionally, I filter out establishments with fewer than one full-time worker. The filtering leaves me with approximately 116 thousand observations of 15 thousand unique establishments.

I utilize information on the value added, wage bill, materials, employment, and investments from the employer data. However, the investment series has some significant outliers, so I apply winsorization to the investments at the one percent level to retain as much information as possible. Then, using the perpetual inventory method, I construct a series for capital stock with the investment data, which evolves according to the formula $k' = (1 - \delta)k + i$, where δ represents the depreciation rate, which I set at 0.1, and i represents the investments. Finally, I define the initial value of capital stock as $k_{\text{first}} = \max\{i_{\text{first}}/\delta, 0\}$.

In the estimation, I use the wage compensations to measure labor input. This choice of labor input measure is a conventional approach to correct the effects of different human capital levels in this type of estimation.[19] To account for differences in workers' skills and focus on the impact of hiring on the establishment's productivity, I adjust the number of hires from high-productivity establishments by multiplying the raw headcount with the hires' average wage divided by the average wage in the manufacturing sector

[18] These include legal forms such as those related to the estate of deceased individuals.

[19] Correcting for the labor inputs' human capital is important as using purely headcount can yield unreasonably large labor input elasticities, as shown in Lochner & Schulz (2024), who make a more serious attempt to measure quality-adjusted labor input.

($h^{+}_{it-1} = hc_{it-1} h^{\star+}_{it-1}$, where * denotes the original head count), resulting in human capital correction $hc_{it-1} = \bar{w^*}_{it-1,\text{hires from better firms}} / \bar{w^*}_{\text{industry}}$. In these calculations, I use the mincer residual as a wage measure w^*, which is obtained from regressing the wage using gender indicators, age, age squared, education level indicators, and year indicators.

Appendix B. Production Function Estimation and Productivity Proxy

The appendix describes the production function estimation in more detail. For convenience, I list below firts the assumptions of the control function approach in ACF:

A1 – Information set: The firm's information set, I_{it}, includes productivity shocks $\{z_{i\tau}\}_{\tau=0}^{t}$. The transitory shocks ϵ_{it} satisfy $E[\epsilon_{it}|I_{it}] = 0$.

A2 – First Order Markov: Productivity shocks evolve according to the distribution $p(z_{it+1}|I_{it}) = p(z_{it+1}|z_{it})$ and the distribution is known to firms and is stochastically increasing in z_{it}.

A3 – Timing: Firms accumulate capital and labor according to functions $\ln(k_{it}) = \kappa(k_{it-1}, i_{it-1})$ and $\ln(n_{it}) = \iota(n_{it-1}, h_{it-1}, s_{it-1})$, where investment i_{it-1} and hiring h_{it-1}, and separations, s_{it-1} are chosen in period $t-1$.[20]

A4 – Scalar Unobservable: Firms intermediate input demand is given by $\ln(m_{it}) = f_t(k_{it}, n_{it}, z_{it})$.

A5 – Strict Monotonicity: $f_t(k_{it}, n_{it}, z_{it})$ is strictly increasing in z_{it}.

In the case of standard approach the goal is to estimate the following set of equations

$$\ln(y_{it}) = \beta_0 + \beta_n \ln(n_{it}) + \beta_k \ln(k_{it}) + z_{it} + \epsilon_{it} \tag{16}$$

$$z_{it} = \Upsilon(z_{it-1}) + \omega_{it}. \tag{17}$$

[20]This assumption is slightly stricter than the original one. However, it will be consistent with the quantitative model as I do not have any period between periods in the model.

To do so, it is often assumed that the output is generated by $y_{it} = \min\{e^{\beta_0+z_{it}+\epsilon_{it}} k_{it}^{\beta_k} n_{it}^{\beta_n}, \beta m_{it}\}$ and then the assumptions imply that the intermediate input function, $\ln(m_{it}) = f(k_{it}, n_{it}, z_{it})$, can be inverted with respect to z_{it} leading to

$$z_{it} = f^{-1}(k_{it}, n_{it}, m_{it}). \tag{18}$$

Substituting the function for productivity into Equation (16) yields,

$$\begin{aligned}\ln(y_{it}) &= \beta_0 + \beta_n \ln(n_{it}) + \beta_k \ln(k_{it}) + f^{-1}(k_{it}, n_{it}, m_{it}) + \epsilon_{it} \\ &= \Psi(k_{it}, n_{it}, m_{it}) + \epsilon_{it},\end{aligned} \tag{19}$$

which can be used to obtain estimate $\hat{\Psi}_t(k_{it}, l_{it}, m_{it})$ either nonparametrically or by using polynomial approximation. Using the estimated $\hat{\Psi}_t$, the estimate for productivity can be obtained

$$\hat{z}_{it}(\beta_n, \beta_k, \beta_0) = \hat{\Psi}(k_{it}, n_{it}, m_{it}) - \beta_0 - \beta_n \ln(n_{it}) - \beta_k \ln(k_{it}), \tag{20}$$

which is conditioned on parameter values β_l, β_k, and β_0. Combining this with the assumed productivity process results in the moment condition

$$\begin{aligned}E[\epsilon_{it} + \omega_{it}|I_{it}] = E[&\ln(y_{it}) - \beta_0 - \beta_n \ln(n_{it}) - \beta_k \ln(k_{it}) \\ &- \Upsilon(\hat{z}_{it-1}(\beta_n, \beta_k, \beta_0))|I_{it}] = 0.\end{aligned} \tag{21}$$

Using a set of instruments from the information set and applying standard GMM estimation methods, the parameters can be estimated. In practice, I use a third-order polynomial to model the control function and the productivity process. Following the approach of De Loecker & Warzynski (2012), I incorporate time dummies when estimating the control function in the first step. Additionally, to significantly reduce computing time, I center the productivity process-related parameters, as described in the appendix of ACF.[21] Also, I include a market share control as suggested in De Ridder et al. (2022).

In the extended version of the estimation, I only change the assumptions

[21]The computational complexity arises in the extended version from the need to recalculate $\hat{h}_{it}$ in each iteration using the mobility-link matrix.

about the productivity process to include the spillover component. This changes the moment condition slightly to

$$
\begin{aligned}
E[\epsilon_{it} + \vartheta_{it}|I_{it}] = E[&\ln(y_{it}) - \beta_0 - \beta_n \ln(n_{it}) - \beta_k \ln(k_{it}) - \Upsilon(\hat{z}_{it-1}(\beta_n, \beta_k, \beta_0)) \\
&- \beta_z \hat{h}^{+}_{it-1}(\hat{z}_{it-1}(\beta_n, \beta_k, \beta_0), \hat{\mathbf{z}}_{t-1}(\beta_n, \beta_k, \beta_0))|I_{it}] = 0. \qquad (22)
\end{aligned}
$$

Otherwise, the method remains unchanged.

Appendix C. Additional Spillover Estimations

Table 7: Correlation between productivity and alternative spillover measure.

	Dependent variable:	
	z_t	
	(1)	(2)
$\frac{\sum^{h_{t-1}} \mathbb{I}(z^s_{t-1} > z^r_{t-1})(z^s_{t-1} - z^r_{t-1})}{n_{t-1}}$	0.121	0.126
	(0.0267)	(0.0268)
3rd ord. polyn. on z_{t-1}	Yes	Yes
hc control	No	Yes

Notes: Robust standard errors in parenthesis. The regression specification follows equation (1), with crucial difference that I use Stoyanov's & Zubanov's (2012) measure of spillovers and, thus, the interpretation of β_z is different to the main text. Here, s and r refer to the sender and receiver, and the indicator function is one if the productivity difference between the sender and receiver is positive. The number of observations is 100,411.

I construct the productivity spillover measure of Stoyanov & Zubanov (2012) to have a connecting point to the empirical literature. Their spillover measure is calculated by taking all the newly hired workers and looking at how the productivity of the new employer and their previous employer deviated from each other in the last period. Then, they calculate the sum of the productivity differences and divide it by employment to get the spillover measure. In Table 7 column 1, I report the results where I regress this alternative measure on productivity, and the results are remarkably similar to those in Stoyanov & Zubanov (2012). Of the results in Stoyanov & Zubanov (2012), the regression in column 2 of Table 14 is the most comparable, and the corresponding multiplier is 0.166, which is larger than the value 0.121 I obtained. The key difference is that I do not include the same control variables, so a direct comparison is not possible. However, I have included a regression in column 2, in which I include the average residual wage of new hires as a control variable, which has no significant influence on the coefficient.

Table 8: Input elasticity estimates and spillover estimates with different specifications.

	Estimation Method:			
	ACF-SO (2*)	ACF-SO (3*)	ACF-SO (4*)	ACF-SO (5*)
n_t	0.700	0.725	0.700	0.700
	[0.680;0.719]	[0.701;0.752]	[0.680;0.719]	[0.679;0.721]
k_t	0.204	0.187	0.204	0.204
	[0.188;0.220]	[0.171;0.216]	[0.186;0.219]	[0.186;0.219]
h^{+}_{t-1}	0.0084			
	[0.0049;0.0143]			
$\frac{h^{+}_{t-1}}{n_{t-1}}$		0.0355		
		[0.0180;0.0562]		
$h^{+}_{t-1}\mathbb{I}(z_{t-1} < \bar{z})$			0.0136	
			[0.0107;0.0198]	
$h^{+}_{t-1}\mathbb{I}(z_{t-1} > \bar{z})$			0.0035	
			[0.0016;0.0101]	
$h^{+\leq 10\%}_{t-1}$				0.0032
				[0.0013;0.0124]
$h^{+>10\%}_{t-1}$				0.0119
				[0.0086;0.0154]
3rd ord. polyn. on z_{t-1}	Yes	Yes	Yes	Yes
hc correction	No	Yes	Yes	Yes

Notes: See notes of Table 2. The column number refers to corresponding column in the Table 1. The first specification drops the hc correction. The second and third column reports the estimates for the relative measure of spillovers (with respect to employment and initial productivity). The final column includes the spillover measure conditional on sending establishments' and receiving establishments' initial productivity differences.

As an additional robustness check, I calculate also the alternative spillover specifications for the consistent approach in Table 8. It contains a robustness check with respect to the human capital adjustment of new hires and the alternative specifications. The estimates of input elasticity remain almost unchanged, as would be expected from the main results. For the spillover effects, we see roughly similar numbers as in the case of the naive approach in the text.

Appendix D. Nature of the Spillovers

The estimation does not determine whether the increase in productivity is due to more efficient workers or general knowledge improvement that leads to higher productivity for all workers. For instance, consider an establishment with two workers having individual productivities h_1, h_2 and a joint establishment productivity component z. The resulting total productivity is $z(h_1 + h_2)$. Now, if h_1 is replaced by a more productive worker $h_3 > h_1$ the productivity would increase. However, we can find a z^* that would result in a similar increase in productivity if $h_1 = h_3$. Even if I use wage stock to measure labor input, which attempts to control worker productivity, and the average wage correction for the hires, some of these concerns remain.

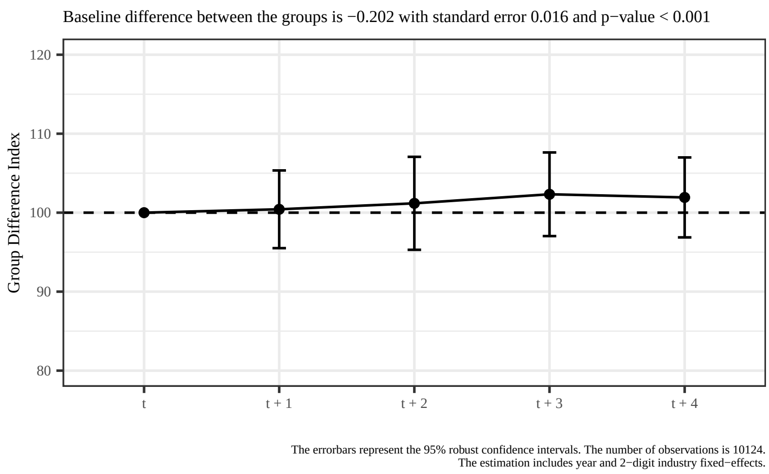


Figure 4: Productivity trends of two groups of establishments that hired workers from a better-performing one. The first group consists of establishments that lost the worker who came from a better-performing unit, while the second group includes establishments from which some other worker left.

To disentangle between these two potential stories, I analyze a specific event. I gather all establishments that hire at least one worker from a more productive unit and calculate their productivity estimates $\hat{z}_{it}$. Then, I separate these establishments based on whether some of the hired workers, who potentially transfer knowledge, leave immediately in the next period. If the productivity evolution significantly differs between these two groups, it provides evidence of worker-specific knowledge. If the productivities evolve similarly, the evidence supports the establishment-specific productivity increase. The exercise is descriptive and helpful in guiding modeling choices.

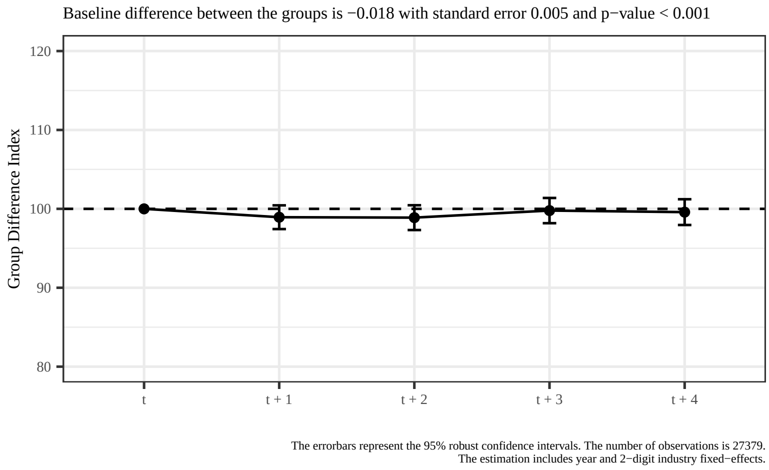


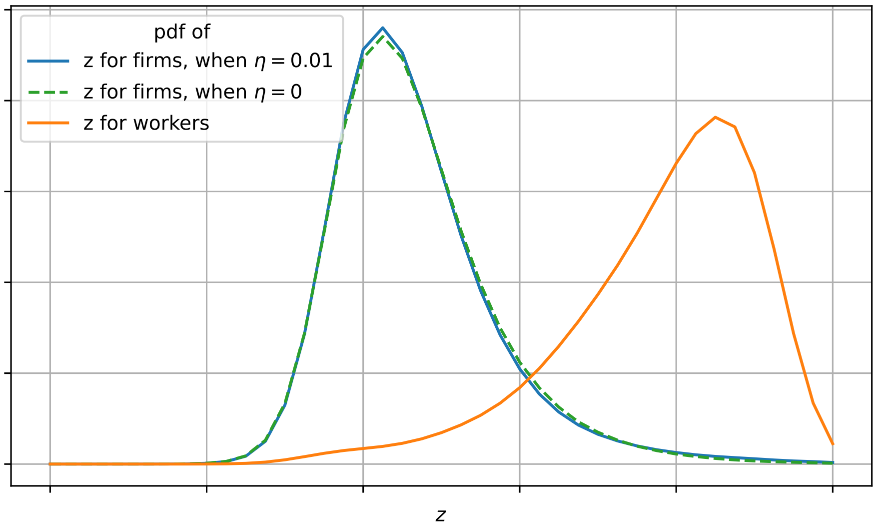
Figure 5: Productivity trends of two groups of establishments that hired workers from a better-performing one. The first group consists of establishments that lost the worker who came from a better-performing unit, while the second group includes rest of the establishments.

Figure 4 demonstrates the productivity changes when a worker hired from a more productive establishment immediately leaves. The difference between the groups is not statistically significant. Thus, the evidence in Figure 4 supports the story of establishment-specific productivity increase through spillovers.

To further examine the robustness of the exercise, I consider different alternative specifications for the control group. For example, I relax the assumption that the ”control” group must have some separations immediately in Figure 5. The results do not indicate a statistically significant difference between the groups. Additionally, based on the first or last applicable event, I check specifications where one establishment occurs only once, either in the ”treated” or ”control” group. These results are not reported here. However, this alternative specification does not affect the result.

In light of this exploration, it seems reasonable to study the effect of spillovers on establishment-specific knowledge as I find no evidence that the spillover effect solely stems from productivity increase through workers' human capital. Therefore, to increase tractability, I ignore the human capital heterogeneity and concentrate on the impact of spillovers on aggregate growth through establishment productivities in the quantitative exercise.

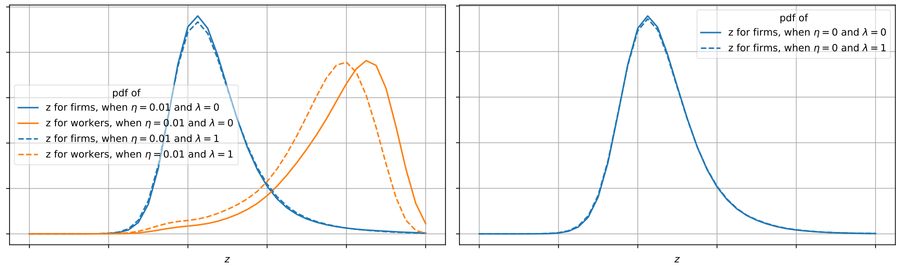
Appendix E. Productivity Distributions



Notes: The figure contains gaussian kernel density estimates for the marginal distribution over z. By using the probability mass function implied by $\mu(\mathbf{x})$, I create simulated data of 100000 observations from the marginal distribution and then fit the kernel to this data. As a bandwith, I use 0.25.

Figure 6: Distributions of establishment and worker productivities.

To illustrate additional central features of the model, I plot the marginal distribution concerning productivity and the workers' knowledge distribution in Figure 6 for the headline results table. We can see that productivity distributions show relatively modest changes. Shutting down the spillover mechanism decreases the number of extremely productive establishments and increases the number of middle-productivity establishments. From the figure, we can also see the central feature of the workers' knowledge distribution, which is that it is heavily centered on large productivities as these establishments are the largest and, when adjusting them, also release relatively more employees.



Notes: The figure contains gaussian kernel density estimates for the marginal distribution over z. By using the probability mass function implied by $\mu(\mathbf{x})$, I create simulated data of 100000 observations from the marginal distribution and then fit the kernel to this data. As a bandwith, I use 0.25.

Figure 7: Distributions of establishment and worker productivities. The left figure represents model with knowledge diffusion when firing costs are introduced. The figure on the right plots the introduction of firing costs without the spillover mechanism.

In Figure 7, I report the changes in the productivity distribution in the firing cost excercise. The figure shows that workers knowledge distribution moves to the left quite dramatically as the size distribution changes. This has consequences for establishment dynamics, which are outlined in the main text.

Appendix F. Calibration of the Model without Spillovers

To recalibrate the model without spillovers, I followed the same procedure as in the benchmark model, with the only difference being that I set the spillover to zero and dropped the corresponding target. The results of the calibration are presented in Table 9. Compared to the benchmark case, we can observe an increase in the entry cost from 3.6 to 6.5, an increase in fixed costs from 0.32 to 0.59, an increase in adjustment costs from 3.2 to 3.6, and a decrease in the entrants' tracking distance from 0.32 to 0.29.

Table 9: Model re-calibration and data fit.

Variable	Value	Explanation	Target	Model
External				
β	0.95	Convention		
α	0.70	ACF estimation		
σ_z	0.31	Std. of entrant prod.		
σ_u	0.14	Std. of prod. diff.		
ψ	0.0	Knowledge mob. adj.		
Internal				
η	0.0	Infered from ACF estimate	0.0056	0.0
f_e	6.5	Infered from mean size	20	20
f_f	0.59	Infered from entry rate	0.067	0.068
f_a	3.6	Infered from job turnover	0.12	0.12
κ	0.29	Infered from growth rate	0.026	0.026