THE UNIVERSITY OF CHICAGO

ESSAYS ON PRODUCTIVITY GROWTH FROM NEW WORKERS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
MICHAEL JAMES KIRKER

CHICAGO, ILLINOIS
JUNE 2019

TABLE OF CONTENTS

LIST OF FIGURES . . . v

LIST OF TABLES . . . vi

1 LEARNING THROUGH HIRING: KNOWLEDGE FROM NEW WORKERS AS AN EXPLANATION OF ENDOGENOUS GROWTH . . . 1
- 1.1 Abstract . . . 1
- 1.2 Introduction . . . 1
- 1.3 Related Literature . . . 3
- 1.4 The Model . . . 5
  - 1.4.1 The Goods Market . . . 6
  - 1.4.2 Wages . . . 8
  - 1.4.3 Worker's Moving Decision . . . 10
  - 1.4.4 The Labor Market . . . 10
  - 1.4.5 The Evolution of Firm Productivity and the Spillover of Knowledge From New Workers . . . 12
  - 1.4.6 Value of a Firm . . . 13
  - 1.4.7 Competitive Equilibrium and Balanced Growth Path of the Model . . . 14
- 1.5 Balanced Growth Path Results . . . 16
  - 1.5.1 Sufficient Conditions for a BGP . . . 17
- 1.6 Conclusions . . . 30

2 FIRM PRODUCTIVITY GROWTH AND ITS RELATIONSHIP TO THE KNOWLEDGE OF NEW WORKERS . . . 32
- 2.1 Abstract . . . 32
- 2.2 Disclaimer . . . 32
- 2.3 Introduction . . . 33
- 2.4 Related Literature . . . 37
- 2.5 The Model . . . 39
  - 2.5.1 Measuring Productivity . . . 39
  - 2.5.2 Modeling Productivity Growth . . . 40
  - 2.5.3 Exposure to Outside Knowledge . . . 43
  - 2.5.4 Theoretical Predictions . . . 47
- 2.6 Data . . . 50
  - 2.6.1 Firm Data . . . 51
  - 2.6.2 Worker Data . . . 52
  - 2.6.3 Summary Statistics . . . 55
- 2.7 Analysis . . . 60
  - 2.7.1 Baseline Model . . . 60
  - 2.7.2 Extensions to the Model . . . 69
  - 2.7.3 Robustness . . . 83
  - 2.7.4 Additional instruments for $\Delta \ln A_{i,j,t-1}$ . . . 85
- 2.8 Conclusions . . . 87

APPENDIX A FURTHER FIRM-LEVEL SUMMARY STATISTICS 90
A.1 Distribution of Firm Productivity 90
A.2 Productivity Dynamics 93
A.3 Distribution of Firm Size (Zipf's Law) 94

APPENDIX B ADDITIONAL REGRESSION SPECIFICATIONS 97
B.1 Alternative Measures of Worker Quality 97
B.2 Average Productivity Differences Between Industries 99
B.3 Stoyanov and Zubanov (2012) Estimation Specification 101
B.4 Controlling for Worker Quality 103

REFERENCES 106

LIST OF FIGURES

A.1 Kernel density estimates of the productivity distributions 91

LIST OF TABLES

2.1 Summary of theoretical predictions 47
2.2 Summary statistics at the firm-year level (Value-added per worker) 56
2.3 Worker transitions — Value-added per worker 57
2.4 Summary statistics for new workers 59
2.5 Initial regression results 61
2.6 Baseline regression results for various productivity measures 66
2.7 Baseline results for capital-labor ratio measure 68
2.8 Regression results featuring between and within industry productivity gaps . . . 72
2.9 Effects of considering worker tenure 76
2.10 Effects of considering worker tenure — capital-labor ratio 79
2.11 Worker flows by worker skill level 82
2.12 Effect of controlling for unobserved productivity shocks 86
2.13 Effect of additional instruments for lagged productivity 88

A.1 Firm productivity transition matrices 93
A.2 Estimates of Zipf's law for the distribution of firm size 95

B.1 Alternative measures of worker quality 98
B.2 Baseline model estimated using demeaned value-added data 100
B.3 Regressions using a functional form similar to Stoyanov and Zubanov (2012) . . 102
B.4 Additional controls for the types on new workers hired 105

CHAPTER 1

LEARNING THROUGH HIRING: KNOWLEDGE FROM NEW WORKERS AS AN EXPLANATION OF ENDOGENOUS GROWTH

1.1 Abstract

This paper develops an endogenous growth model in which the job-to-job transition of workers provides a channel for the spillover of knowledge between firms. Workers learn some of the productive knowledge used by their employer while working on the job. When a worker moves to another firm, they are able to adapt some of this knowledge for use at the hiring firm. Firms endogenously control their exposure to new knowledge by choosing the intensity that they post vacancies in a search-and-matching labor market. It is shown that under a set of assumptions regarding the initial distribution of firm types and the vacancy posting cost function, the competitive equilibrium leads to a balanced growth path that has a constant growth rate and stationary distribution of firm size.

1.2 Introduction

Since the seminal work on endogenous growth by Romer (1986) and Lucas (1988), the spillovers (or positive externalities) of new knowledge have been viewed as an important dynamic in generating sustainable long-run growth.[1] One channel through which knowledge is thought to spillover between firms is through hiring workers with experience at other firms. Workers are exposed to their firm's production, managerial, and marketing knowledge on a daily basis. While some of this knowledge can be protected (e.g. patents) not all knowledge can. Therefore, when workers moves to another firm they are able to adapt some of the knowledge obtained at their previous employer for use at their new job, improving the stock

1. See Griliches (1992) for an early review of R&D spillovers.

of productive knowledge used at the hiring firm. What's more, incumbent workers at the hiring firm will be able to absorb some of the information that new worker brings, and can diffuse the knowledge further if they switch jobs, multiplying the initial social impact of the knowledge diffusion.

There is empirical support for the notion that firms learn from the new workers they hire. According to the 2013 Business Operations Survey, a nationally representative survey of New Zealand firms by Statistics New Zealand (the national statistical agency), 52 percent of businesses who reported undertaking some form of innovation in the previous two years stated that new staff were an important source of ideas for the innovation.[2] Furthermore, analysis of linked employer-employee data by Stoyanov and Zubanov (2012), Parrotta and Pozzoli (2012), and Serafinelli (2015) find correlations between firm-level hiring patterns and productivity growth that support the idea of knowledge spillover from new hires.

This paper develops an endogenous growth model in which learning from new hires is the mechanism by which knowledge diffuses between firms. Workers are assumed to passively absorb the productive knowledge of their current employer. They carry out on-the-job searching for new employment opportunities, and when they move to a less productive firms, they are able to transplant some of their knowledge into the hiring firm. The firms themselves are heterogeneous in how many workers they employ and the stock of productive knowledge they use. A firm's exposure to new knowledge is endogenously determined by the firm.

Standard endogenous growth models that feature knowledge spillover typically abstract from the mechanism by which firms learn, and usually assume an exogenous learning rate.

2. The types of innovation asked about were: (i) product innovation: "did this business introduce into the market any new or significantly improved goods or services?"; (ii) process innovation: "did this business implement any new or significantly improved operational processes (i.e. methods of producing or distributing goods or services)?"; (iii) organizational innovation: "did this business implement any new or significantly improved organizational/managerial processes (i.e. significant changes in this businesses strategies, structures or routines)?"; and (iv) marketing innovation: "did this business implement any new or significantly improved sales or marketing methods which were intended to increase the appeal of goods or services for specific market segments or to gain entry to new markets?"

The novel feature of this model is that it gives structure to the learning mechanism. Within the model of this paper, firms optimally choose the rate at which they learn by varying their vacancy posting rate. A higher vacancy posting rate leads to a greater inflow of new workers and hence a greater exposure to new knowledge, but also incurs a greater search cost. In addition, the rate at which vacancies posted by any firm are matched to new workers is endogenously determined by the tightness in the labor market.

The aim of this paper is to show that there is a certain set of assumptions that are sufficient for the competitive equilibrium to be a balanced growth path. Along this balanced growth path all firms improve productivity at the same rate and the distribution of firm size is constant throughout time. Such a balanced growth path is consistent with those in standard endogenous growth models featuring knowledge spillover at an exogenous rate. As a result, learning from new hires can be viewed as one possible explanation for how firms learn from other firms in standard endogenous growth models.

The rest of the paper is organized as follows. Section 1.3 discusses how this paper fits into the existing literature. The structure of the theoretical model is presented in Section 1.4. Section 1.5 presents a set of assumption on the initial distribution of firm type and the vacancy posting cost function and shows that these assumptions are sufficient to generate a balanced growth path. Section 1.6 concludes.

1.3 Related Literature

Standard endogenous growth models that rely on the spillover of knowledge typically assume that agents receive learning opportunities costlessly at an exogenous rate (see Luttmer, 2012 and Luttmer, 2015 as examples). However, some progress has been made towards adding structure to the learning process. Perla and Tonetti (2014) and Lucas and Moll (2014) develop models that introduce an endogenous search effort decision into the knowledge spillover framework. Within both models, agents optimally trade off the amount of time they spend spend producing today with the time they spend searching for new learning opportunities to become more productive in the future.

Similar to those two models, the model developed in this paper features an endogenous search effort choice embedded within the firm's vacancy posting rate decision. Firms that post more vacancies attract more workers and are therefore exposed to more knowledge spillover from new workers. However, they also incur greater search costs. Furthermore, the rate at which firms learn depends not only on their choice of vacancy postings, but also the endogenous tightness of the labor market which is related to the vacancy posting rate of all other firms.

Previous papers that have studied labor mobility as a channel for diffusing knowledge in a dynamic context primarily have used an overlapping-generations framework.[3] For example, Dasgupta (2011) and Monge-Naranjo (2012) both develop overlapping-generations models in which workers learn on the job when young and use the knowledge that they acquire to start their own firms later in life. Within these types of model, knowledge spillover is only from incumbent firms to new firms. The model developed in this paper uses the hiring of new workers as a way to explain how incumbent firms are able to learn from other incumbent firms. It also includes a search-and-matching labor market for the workers which introduces frictions into the knowledge spillover process.

The labor market used in this model draws on the modeling techniques developed in the search-and-matching literature. On one side of the labor market workers search while on the job, as in the model developed by Mortensen (2010). On the other side of the labor market, large (employing multiple workers) heterogeneous firms post job vacancies similar to the models of Mortensen (2010) and Acemoglu and Hawkins (2014). The wage setting equation used in this paper is also adopted from the approach taken in both models.

3. Partial equilibrium analysis of this same channel has also been conducted in the area of game theory. For example see Fosfuri et al. (1998) and Glass and Saggi (2002).

1.4 The Model

Time is continuous. There exists a continuum of firms of measure one, and continuum of workers of measure one. Both firms and workers are infinitely lived. Firms are large (each employing a measure of workers), have different stocks of productive knowledge, and produce differentiated goods. The only active decision made by firms is the rate at which to post job vacancies in order to hire new workers. Hiring new workers benefits the firm in two ways. First, it increases the amount of labor the firm can utilize in the production process. Second, it increases the stock of productive knowledge at the firm due to knowledge spillovers from workers with previous experience at firms with superior productive knowledge.

All workers supply one unit of labor inelastically. Where workers differ is in their level of knowledge. Through their participation in the production process, workers absorb the productive knowledge of their current employer. Workers who move to a less productive firm are capable of adapting some of this knowledge for use at the hiring firm.

For simplicity, the model abstracts from both the labor-market search intensity decision of workers, and the negotiated premium firms would be willing to pay worker to entice them away from their previous employer. Instead, all workers are assumed to search for new jobs with the same search effort, and all workers will choose to move to a new firm if the wage paid by the hiring firm is at least as high as the wage paid by the worker's current employer. These simplifications remove the need to explicitly model the value of the worker.

Below, the details of the goods market, labor market, knowledge diffusion technology, and the decision choice of the firms are presented.

1.4.1 The Goods Market

The Demand for the Final Consumption Good

Each worker belongs to an identical household. The households have an inter-temporal utility function given by

$$\mathbb{E}_0\left[\int_0^\infty e^{-\kappa t}U(C(t))\,dt\right] \;=\; \mathbb{E}_0\left[\int_0^\infty e^{-\kappa t}\frac{C(t)^{1-\sigma}}{1-\sigma}\,dt\right],$$

where $\kappa > 0$ is the rate of time preference, $1/\sigma > 0$ is the elasticity of inter-temporal substitution, $U(\cdot)$ is the period utility function, and $C(t)$ is the amount of final consumption good consumed at date t.

The final consumption good comprises of a Dixit-Stigliz CES aggregation of the individual goods produced by the continuum of firms,

$$C(t) = \left[\int_{i=0}^{1} c(i,t)^{(\rho-1)/\rho}\,di\right]^{\rho/(\rho-1)}, \qquad \text{all } t,$$

where $c(i,t)$ is the consumption of the good produced by firm $i \in [0,1]$, and $\rho > 1$ is the elasticity of substitution between the output of different firms.

From the household's optimization problem, the inverse demand for the output of firm i is given by,

$$p(i,t) = P(t)\left(\frac{c(i,t)}{C(t)}\right)^{-1/\rho}, \qquad \text{all } i \in [0,1],\ t, \tag{1.1}$$

where $p(i,t)$ is the price of good i and

$$P(t) = \left[\int_{i=0}^{1} p(i,t)^{1-\rho}\,di\right]^{1/(1-\rho)}, \qquad \text{all } t,$$

is the aggregate price index at date t. Without loss of generality, prices will be normalized so that $P(t) = 1$ for all t.

The household's preferences also determine the real interest rate,

$$r(t) = \kappa + \sigma g(t), \qquad \text{all } t, \tag{1.2}$$

where

$$g(t) = \frac{d}{dt} \ln C(t), \qquad \text{all } t, \tag{1.3}$$

is the growth rate of the final consumption good.

Firms

Each firm produces a unique variety of good and is monopolistically competitive in the goods market. Firms are also heterogeneous along two other dimensions. First, each firm has a stock of productive knowledge that acts as technology in the production function. The minimum level of knowledge/productivity at any firm at date $t = 0$ is $z_{min,0} \geq 0$. Second, firms differ by the labor employed, $l \in \mathbb{R}_{\geq 0}$. Throughout this paper, the term "firm size" will always refer to the amount of employment at a firm.

Let $\mathcal{S}$ denote the joint productivity-size space for firms, and let $s \equiv (z, l)$ identify a state in $\mathcal{S}$ for a firm. The state of a particular firm will be referred to as the firm's "type". The distribution of firm types at date t is denoted by $F(s, t)$.

All firms combine their stock of productive knowledge with the labor of their workers to produce their unique good using the following production function

$$y(s) = z l^{\alpha}, \qquad \text{all } s \in \mathcal{S}, \tag{1.4}$$

where $0 < \alpha \leq 1$ is the labor elasticity parameter.

Because the output of all firms enter the CES aggregator for the final consumption good symmetrically, firms can be indexed by their type (s) rather than the good they produce (i). Therefore, aggregate output of the final good at date t can be written as

$$Y(t) = \left[\int_{\mathcal{S}} y(s)^{(\rho-1)/\rho} \, dF(s,t) \right]^{\rho/(\rho-1)}, \qquad \text{all } t. \tag{1.5}$$

Given the demand that firms faces, the revenue of an s-type firm operating at date t is

$$R(s,t) = p(s,t)y(s), \qquad \text{all } s \in \mathcal{S},\, t. \tag{1.6}$$

1.4.2 Wages

The wage setting mechanism between a firm and its workers follows the approach used by both Mortensen (2010) and Acemoglu and Hawkins (2014). It is a generalization of Nash-bargaining for one large firm negotiating with many workers, in the spirit of Stole and Zwiebel (1996).

Long term wage contracts are not feasible. Wage negotiations are costless and instantaneous, allowing the firm to renegotiate individually with all workers in as many negotiation rounds as needed before production takes place at each point in time.

If the firm and a particular worker are not able to agree to a mutually satisfactory wage, the worker sits out the production process and earns zero income at that date. The firm will then produce output using the pool of workers that it was able to negotiate a mutually satisfactory wage with. Because all workers at the firm are homogeneous in the amount of labor they supply and have the same bargaining strength, in equilibrium all workers will agree to work, and the wage for all workers at the firm, $\omega(s,t)$, will satisfy a surplus sharing rule

$$\beta \frac{\partial \pi(s,t)}{\partial l} = (1-\beta)\omega(s,t), \qquad \text{all } s \in \mathcal{S},\, t, \tag{1.7}$$

where $\partial \pi(s,t)/\partial l$ denotes the firm's marginal profit of labor, $0 < \beta < 1$ denotes the worker's relative bargaining strength, and the wage $\omega(s,t)$ measures the worker's marginal benefit from working.

The firm's profit function, $\pi(s,t)$, has the standard form,

$$\pi(s,t) \quad = \quad R(s,t) - \omega(s,t)l, \qquad \text{all } s \in \mathcal{S},\, t, \tag{1.8}$$

where $\omega(s,t)l$ is the total wage cost to the firm for all units of labor employed.

Substituting (1.8) and (1.6) into (1.7) and solving for the wage rate ω shows that the wage at an s-type firm at date t is equal to a β share of the firm's marginal revenue of labor,

$$\omega(s,t) = \beta \frac{\partial R(s,t)}{\partial l}, \qquad \text{all } s \in \mathcal{S},\, t. \tag{1.9}$$

Substituting (1.9) and (1.1) into (1.8) yields the profit function of the firm as a function of the firm's output,

$$\widehat{\pi}(y(s),t) = (1-\widehat{\beta})P(t)Y(t)^{1/\rho}y(s)^{1-1/\rho}, \qquad \text{all } s \in \mathcal{S},\, t, \tag{1.10}$$

where $\widehat{\beta} \equiv \beta\alpha(1-1/\rho)$.

Similarly, substituting (1.6) and (1.1) into (1.9) provides the wage as a function of the firm's output,

$$\widehat{\omega}(y(s),t) = \beta P(t)Y(t)^{1/\rho}\frac{\partial y(s)^{1-1/\rho}}{\partial l}, \qquad \text{all } s \in \mathcal{S},\, t. \tag{1.11}$$

The average wage across all workers at date t is defined as

$$\overline{\omega}(t) = \int_{\mathcal{S}} l'\omega(z',l',t)\, dF(s',t), \qquad \text{all } t. \tag{1.12}$$

1.4.3 Worker's Moving Decision

Workers seek to maximize their current wage. Therefore, when a worker is matched with another firm in the labor market, the worker will choose to change firms if the wage offered by the outside firm is equal or greater than the wage paid by their current employer.

Let $\mathcal{A}(s,t)$ be the set of type $s' = (z', l')$ firms that an $s = (z, l)$ type firm can poach workers from (i.e. workers at s'-type firms would choose to move to an s-type firm if matched). The region $\mathcal{A}(s,t)$ is defined by the set of type s' firms where

$$\mathcal{A}(s,t) \equiv \left\{ s' \in \mathcal{S} \,:\, \omega(s,t) - \omega(s',t) \geq 0 \right\}, \qquad \text{all } s \in \mathcal{S},\, t. \tag{1.13}$$

Furthermore, let $\widehat{\mathcal{A}}(s,t)$ denote the sub-region of $\mathcal{A}(s,t)$ where $z' > z$ (i.e. the set of more productive firms from which an $s = (z, l)$ type firm can hire from),

$$\widehat{\mathcal{A}}(s,t) \equiv \left\{ s' \in \mathcal{S} \,:\, \omega(s,t) - \omega(s',t) \geq 0 \cap z' > z \right\}, \qquad \text{all } s \in \mathcal{S},\, t. \tag{1.14}$$

Similarly, let $\mathcal{B}(s,t)$ be the set of type s' firms that can poach workers from an s-type firm at time t,

$$\mathcal{B}(s,t) \equiv \left\{ s' \in \mathcal{S} \,:\, \omega(s',t) - \omega(s,t) \geq 0 \right\}, \qquad \text{all } s \in \mathcal{S},\, t. \tag{1.15}$$

1.4.4 The Labor Market

The labor market is modeled as a single, frictional, search-and-matching market featuring large firms and workers who search while on the job. For simplicity there is no unemployment state for workers.

The Firm's Side of the Labor Market

At each date, firms can choose the rate at which to post job vacancies for new workers. Denote the measure of vacancies an s-type firm posts as $n(s,t) \geq 0$. Firms are assumed to incur a real cost for posting vacancies denoted as $c(n,s,t)$.

The Worker's Side of the Labor Market

For simplicity, this paper abstracts from the search effort decision of workers. Instead, all workers are assumed to search on the job with unit intensity at all points in time. Any search costs to the worker are sunk costs that are constant across all workers, and therefore do not affect the the worker's decisions in any way.

Labor-Market Matching Technology

Posted vacancies and searching workers are matched at random. The measure of matches is given by a homogeneous of degree one Cobb-Douglas matching function. Following the notation of Mortensen (2010), it is convenient to define $\theta(t)$ as the measure of tightness in the labor market (from the firm's perspective) at date t. Because there is a unit mass of workers and each worker searches with an intensity of unity, the total search effort by workers is unity. As a result, $\theta(t)$ can be simplified to the average vacancy postings rate at date t,

$$\theta(t) = \int_{\mathcal{S}} n(s,t)\, dF(s,t), \qquad \text{all } t. \tag{1.16}$$

The probability that a firm's posted job vacancy is matched with a searching worker is given by the matching function

$$q(\theta) \equiv \overline{q}\theta^{-\mu}, \tag{1.17}$$

where $0 < \mu < 1$ is the matching elasticity, and $\overline{q} > 0$ is a normalizing scalar which can either be interpreted as capturing the efficiency of the matching function or as an exogenous (and constant) probability that an acceptable match results in the worker moving.

From the point of view of a firm, posted vacancies are matched at the Poisson rate $n(s,t)q(\theta(t))$. Because workers are infinitesimally small in size, the continuum of posted vacancies by the firm are matched in the labor market with a continuum of searching workers drawn from the entire distribution of searching workers, $lF(s',t)$. From the perspective of workers, their search effort yields a matches at the Poisson rate $\theta(t)q(\theta(t))$.

1.4.5 The Evolution of Firm Productivity and the Spillover of Knowledge From New Workers

Productive knowledge is equally applicable to all firms. Therefore a firm's productivity level, z, can be used to rank firms by their knowledge. When workers move between firms, they take with them their productive knowledge from their previous employer to their new employer. Only those workers hired from more productive firms contribute to the learning of new knowledge at the hiring firm. Workers from less productive firms have inferior knowledge that the hiring firm will choose to disregard in favor of their current productive practices. Therefore, when workers move to more productive firms it is the worker who learns from the firm, and the worker can potential diffuse this new knowledge in a subsequent job move.

When a worker is matched with another firm, the probability that a worker can successfully transfer knowledge from their current employer to the new firm is τ/l, where $0 < \tau \leq 1$ is a reduced-form parameter reflecting the efficiency at which knowledge can be transferred between firms. The term $1/l$ reflects the fact that workers from larger firms are less likely to diffuse productivity knowledge.

At each point in time a firm hires a continuum of workers from other firms. Each of these workers brings with them some new knowledge the hiring firm can exploit. This type of learning process is what Buera and Lucas (2018) refer to as "continuous arrival" of ideas. This represents the limiting case where each individual worker has a chance to spill over knowledge at a Poisson rate. As the size of each worker goes to zero, the number of learning opportunities increases for the firm (as the firm is hiring more workers), but this benefit is offset by a reduction in the ability of each individual worker to transfer knowledge.

Formally, the productivity level of an $s = (z, l)$ type firm evolves as

$$\frac{dz}{dt} = \tau n q(\theta(t)) z \int_{\widehat{\mathcal{A}}(s,t)} \ln(z'/z)\, dF(s', t), \qquad \text{all } z \geq z_{min,0},\ t, \tag{1.18}$$

where n is the number of vacancies posted which yields $nq(\theta(t))$ total matches in the labor market, and the expected improvement in productivity level from a match is given by $z\tau \int_{\widehat{\mathcal{A}}(s,t)} \ln(z'/z)\, dF(s', t)$ which depends in part upon the set of workers with superior knowledge that the firm can successfully poach, $\widehat{\mathcal{A}}(s, t)$ as defined in (1.14).

1.4.6 Value of a Firm

The value of any $s = (z, l)$ type firm at date t, denoted by $\Pi(s, t)$, satisfies the following Hamilton-Jacobi-Bellman equation

$$r(t)\Pi(s,t) = \pi(s,t) + \max_{n \geq 0} \left\{ -c(n, s, t) + \frac{\partial \Pi(s,t)}{\partial z}\frac{dz}{dt} + \frac{\partial \Pi(s,t)}{\partial l}\frac{dl}{dt} + \frac{\partial \Pi(s,t)}{\partial t} \right\}, \text{all } s \in \mathcal{S},\ t, \tag{1.19}$$

where $\pi(s, t) - c(n, s, t)$ is the flow of profits net of vacancy posting costs, the firm's stock of productive knowledge evolves over time according to (1.18), and the size of the firm evolves over time according to

$$\frac{dl}{dt} = nq(\theta(t)) \int_{\mathcal{A}(s,t)} l'\, dF(s', t) - l\theta(t)q(\theta(t)) \int_{\mathcal{B}(s,t)} dF(s', t), \qquad \text{all } l,\ t. \tag{1.20}$$

The first component on the right hand side (RHS) of (1.20) represents the inflow of new labor that the firm successfully poaches from other firms belonging to the set $\mathcal{A}(s, t)$. The second component represents the outflow of incumbent workers who are poached by other firms belonging to the set $\mathcal{B}(s,t)$.

The firm's optimal choice of vacancy postings is denoted by

$$\nu(s,t) = \arg\max_{n\geq 0}\left\{-c(n,s,t) + \frac{\partial \Pi(s,t)}{\partial z}\frac{dz}{dt} + \frac{\partial \Pi(s,t)}{\partial l}\frac{dl}{dt} + \frac{\partial \Pi(s,t)}{\partial t}\right\},$$

and will satisfy the FOC:

$$\begin{aligned}\frac{\partial c(\nu(s,t),s,t)}{\partial \nu(s,t)} =& \frac{\partial \Pi(s,t)}{\partial z}\tau q(\theta(t)) z \int_{\widehat{\mathcal{A}}(s,t)} \ln(z'/z)\, dF(s',t) \\ &+ \frac{\partial \Pi(s,t)}{\partial l} q(\theta(t)) \int_{\mathcal{A}(s,t)} l'\, dF(s',t), \qquad \text{all } n \geq 0,\ s \in \mathcal{S},\ t, \end{aligned} \tag{1.21}$$

which equates the marginal cost of posting vacancies (the left hand side, LHS) with the marginal benefit of a vacancy (the RHS). The marginal benefit comprises of two terms, the marginal benefit from knowledge spillover, and the marginal benefit of additional labor.

1.4.7 Competitive Equilibrium and Balanced Growth Path of the Model

This section presents the definitions of a Competitive Equilibrium (CE) and Balanced Growth Path (BGP) for the model. The main focus of the paper is to examine a set of assumptions under which the CE of the model leads to a BGP along which productivity grows at a constant rate and the distribution of firm size is constant.

Definition 1. Given the functions $[y(s),\ c(\nu,s,t),\ q(\theta)]$, parameters $[\alpha,\ \rho,\ \beta,\ \overline{q},\ \mu,\ \kappa,\ \sigma,\ \tau]$, and initial distribution function $F(s,0)$, the aggregate functions $\{Y(t),\ g(t),\ r(t),\ \theta(t)$, for all $t\geq 0\}$, individual functions $\{p(s,t),\ R(s,t),\ \pi(s,t),\ \omega(s,t),\ \Pi(s,t),\ \nu(s,t)$, for all $s\in\mathcal{S}$ and $t\geq 0\}$, correspondence $\{\mathcal{A},\ \mathcal{B},\ \widehat{\mathcal{A}}$, for all $s\in\mathcal{S}$ and $t\geq 0\}$, and distribution functions $F(s,t)$ for $t>0$, are a *competitive equilibrium* (CE) if

1. $Y(t)$, $p(s,t)$, $R(s,t)$, $g(t)$, and $r(t)$ satisfy (1.1)-(1.3), (1.5) and (1.6);

2. $\omega(s,t)$ and $\pi(s,t)$ satisfy (1.7) and (1.8);

3. $\mathcal{A}(s,t)$, $\widehat{\mathcal{A}}(s,t)$, and $\mathcal{B}(s,t)$ satisfy (1.13) to (1.15);

4. $\theta(t)$ satisfies (1.16);

5. $\Pi(s,t)$ satisfies (1.18) to (1.20), and $\nu(s,t)$ is the maximizing value for n;

6. The evolution of $F(\cdot,t)$ is consistent with both (1.18) and (1.20), and

$$\begin{aligned}\int_{\mathcal{S}} \tilde{l}\, dF(\tilde{z},\tilde{l},t) =& 1,\\ \int_{\mathcal{S}} dF(\tilde{z},\tilde{l},t) =& 1, \qquad \text{all } t.\end{aligned}$$

The two conditions in the final requirement above state that at each date t, total employment is unity, and the total mass of firms is also unity. Therefore, $F(s,t)$ is a proper density function at every date.

Characterizing a competitive equilibrium involves solving a fixed point problem in the vacancy posting function $\nu(s,t)$. The vacancy posting choice of any individual firm must maximize the value of that firm taking as given the policy rules followed by all other firms.

The BGP can be defined as follows:

Definition 2. A competitive equilibrium is a *balanced growth path* if

$$F(z,l,t) = F(e^{-\gamma t}z,l,0) \equiv F_0(e^{-\gamma t}z,l), \qquad \text{all } z \geq z_{min,0},\ l \geq 0,\ t.$$

On a BGP, aggregate output grows at a constant rate γ, and the interest rate is constant,

$$\begin{aligned}Y(t) =& e^{\gamma t}Y_0,\\ g(t) =& \gamma,\\ r(t) =& \overline{r} \equiv \eta + \sigma\gamma, \qquad \text{all } t,\end{aligned}$$

where Y_0 is the aggregate output at date $t = 0$.

The labor market tightness is constant on a BGP,

$$\theta(t) = \overline{\theta}, \qquad \text{all } t.$$

In addition, firm-level prices, revenues, profits, and wages on a BGP satisfy

$$\begin{aligned}
p(z, l, t) &= p(e^{-\gamma t} z, l, 0) \equiv p_0(e^{-\gamma t} z, l), \\
R(z, l, t) &= e^{\gamma t} R(e^{-\gamma t} z, l, 0) \equiv e^{\gamma t} R_0(e^{-\gamma t} z, l), \\
\pi(z, l, t) &= e^{\gamma t} \pi(e^{-\gamma t} z, l, 0) \equiv e^{\gamma t} \pi_0(e^{-\gamma t} z, l), \\
\omega(z, l, t) &= e^{\gamma t} \omega(e^{-\gamma t} z, l, 0) \equiv e^{\gamma t} \omega_0(e^{-\gamma t} z, l), \qquad \text{all } z \geq z_{min,0},\ l \geq 0,\ t.
\end{aligned}$$

As a result, the average wage rate on the BGP grows at the rate γ,

$$\overline{\omega}(t) = e^{\gamma t} \overline{\omega}(0) \equiv e^{\gamma t} \overline{\omega}_0, \qquad \text{all } t.$$

Hence the correspondences describing where firms can poach and be poached from satisfy

$$\begin{aligned}
\mathcal{A}(z, l, t) &= \mathcal{A}(e^{-\gamma t} z, l, 0) \equiv \mathcal{A}_0(e^{-\gamma t} z, l), \\
\widehat{\mathcal{A}}(z, l, t) &= \widehat{\mathcal{A}}(e^{-\gamma t} z, l, 0) \equiv \widehat{\mathcal{A}}_0(e^{-\gamma t} z, l), \\
\mathcal{B}(z, l, t) &= \mathcal{B}(e^{-\gamma t} z, l, 0) \equiv \mathcal{B}_0(e^{-\gamma t} z, l), \qquad \text{all } z \geq z_{min,0},\ l \geq 0,\ t. \qquad (1.22)
\end{aligned}$$

1.5 Balanced Growth Path Results

The main goal of this paper is to show that under a certain set of assumption, including a linear form for $c(\cdot)$, there exists a BGP. Attention is focused on the BGP of the model where the joint distribution of firm types at date $t = 0$ has the property that all firms with the same level of productive knowledge also have the same firm size. This type of distribution will be referred to as a *one-dimensional distribution*.

Define $\Phi(z,t)$ as the marginal distribution function for z under F at date t,

$$\Phi(z,t) = \int_{z_{min,0}}^{z} \int_{0}^{\infty} f(\zeta, l, t)\, dl\, d\zeta, \qquad \text{all } z \geq z_{min,0},\ t.$$

Using Φ, a one-dimensional distribution can be formally defined as follows:

Definition 3. A distribution $F(\cdot, t)$ is *one-dimensional* if, for some function $\psi(z)$,

$$\Phi(z,t) = \lim_{\varepsilon \to 0^+} \int_{z_{min,0}}^{z} \int_{-\varepsilon}^{+\varepsilon} f(\zeta, \psi(\zeta) + u)\, du\, d\zeta, \qquad \text{all } z \geq z_{min,0},\ t.$$

Therefore, in a one-dimensional distribution all of the distribution's mass lies on the manifold $[z, \psi(z)]$. If the initial distribution is one-dimensional, then clearly it retains that property on a BGP,

$$\Phi(z,t) = \Phi(e^{-\gamma t} z, 0) \equiv \Phi_0(z e^{-\gamma t}), \qquad \text{all } z \geq z_{min,0},\ t.$$

1.5.1 Sufficient Conditions for a BGP

This section details a set of assumptions that are sufficient to generate a BGP when the initial distribution is one-dimensional. The first two assumption relates to the initial distribution of firm type.

Assumption 1. *The initial marginal distribution function for z is a Pareto distribution with location and shape parameters $(z_{min,0}, \chi)$, where*

$$\chi \equiv \frac{1 - 1/\rho}{1 - \alpha(1 - 1/\rho)}.$$

Assumption 2. *The initial distribution of firm size satisfies*

$$\psi(z) =\psi_0 z^{\chi}, \qquad z \geq z_{min,0}, \tag{1.23}$$

where ψ_0 is a constant, determined by the fact that aggregate employment is unity,

$$1 =\psi_0 \int_{z_{min,0}}^{\infty} z^{\chi} \, d\Phi_0(z).$$

According to Assumption 2, the relative size of any two firms is a function of the knowledge/productivity difference between the firms. This leads to the following property in regards to the wage at each firm.

Lemma 1. *Under Assumption 2 the wage rate is the same across all firm types at date $t = 0$.*

Proof. Using (1.11), (1.4), and $P(t) = 1$, the wage rates at date $t = 0$ are

$$w_0(z) = W_0 z^{1-1/\rho} \psi(z)^{\alpha(1-1/\rho)-1}, \qquad \text{all } z \geq z_{min,0},$$

where W_0 is a constant that is the same across all firm types.

Assumption 2 implies that

$$z^{1-1/\rho} \psi(z)^{\alpha(1-1/\rho)-1} = (z')^{1-1/\rho} \psi(z')^{\alpha(1-1/\rho)-1}, \qquad \text{all } z \geq z_{min,0},\ z' \geq z_{min,0},$$

which when combined with the expression for $w_0(z)$ above shows that the wage rate is the same at all firms. Q.E.D.

The final assumption made relates to the properties of the vacancy posting cost function.

Assumption 3. *The vacancy posting cost function, $c(n,s,t)$, has the form*

$$c(n,s,t) = l\widehat{c}(n/l)\overline{w}(t), \qquad \text{all } n \geq 0,\ s \in \mathcal{S},\ t, \tag{1.24}$$

where $\widehat{c}(\cdot)$ is a strictly increasing, strictly convex, and differentiable function with $\widehat{c}(0) = \widehat{c}'(0) = 0$, and $\overline{w}(t)$ is the economy-wide average wage rate at date t as previously defined in (1.12).

Under Assumption 3, the vacancy posting cost does not depend upon the firm's stock of productive knowledge. In addition, $c(\cdot)$ is homogeneous of degree one in (n,l), and along the BGP the vacancy posting cost grows at the constant rate γ.

Having outlined the key assumptions, it is now possible to state the main proposition for a BGP.

Proposition 1. *If the initial distribution $F(s,0)$ satisfies Assumptions 1 and 2, and the vacancy posting cost function satisfies Assumption 3, then the competitive equilibrium is a balanced growth path.*

In order to prove Proposition 1 several properties of firms along the one-dimensional manifold of the BGP will first be established in a series of Lemmas. These Lemmas make use of the assumptions given in Proposition 1. Lemma 2 establishes all the possible moves of workers between firms on the BGP. Lemma 3 establishes the vacancy posting rate necessary for firms on the BGP to maintain their size at all dates. Lemma 4 establishes the value of firms that are on the one-dimensional manifold of the BGP and shows that the vacancy posting rate necessary for firms to maintain their size on the BGP is also the rate that causes the productivity of each firm to grow at the same constant rate.

Using the details from these Lemmas, the proof of the main proposition then proceeds in two parts. It is first shown that in a competitive equilibrium, all firm types will choose the same ratio of vacancy postings to firm size, and this rate coincides with the one shown in the Lemmas to maintain each firm's size along the BGP. It is then shown that when all firms are following this vacancy posting rule, the value of the labor market tightness in the competitive equilibrium is unique in value.

Lemma 2. *On the BGP the correspondences $\mathcal{A}$, $\widehat{\mathcal{A}}$, and $\mathcal{B}$ in (1.22) satisfy*

$$\begin{aligned}
&\mathcal{A}(z,\psi(z),0) \supseteq \left\{s' \in \mathcal{S} : s' = (z',\psi(z'))\right\}, \\
&\mathcal{B}(z,\psi(z),0) \supseteq \left\{s' \in \mathcal{S} : s' = (z',\psi(z'))\right\}, \\
&\widehat{\mathcal{A}}(z,\psi(z),0) = \left\{s' \in \mathcal{A}(z,\psi(z),0) \text{ and } z' > z\right\}, \qquad \text{all } z \geq z_{min,0}.
\end{aligned}$$

Proof. Immediately follows from (1.22) and Lemma 1. Q.E.D.

The next result, Lemma 3, shows that employment remains constant for every firm on the one-dimensional manifold if, and only if, all firms post vacancies at the rate $n(z,l) = \overline{\theta} l$, where $\overline{\theta}$ is the labor market tightness.

Lemma 3. *On a BGP, if one exists, every firm on the one-dimensional manifold must choose a vacancy posting rate that is proportional to its current employment $\nu(z,l) = \overline{\theta} l$, where $\overline{\theta}$ is the labor market tightness.*

Proof. Using (1.20) and Lemma 2, for a firm of initial type $(\zeta, l) = [\zeta, \psi(\zeta)]$, employment is constant over time if and only if

$$0 = \frac{dl}{dt} = \nu\left(e^{\gamma t}\zeta, \psi(\zeta)\right) \int_{z_{min,0}}^{\infty} \psi(\zeta')\, d\Phi_0(\zeta') - \overline{\theta}\psi(\zeta), \qquad \text{all } t, \tag{1.25}$$

where $\int_{z_{min,0}}^{\infty} \psi(\zeta')\, d\Phi_0(\zeta') = 1$ by the fact that there is measure one of workers.

From before, (1.25) implies that, given $\overline{\theta}$, employment remains constant at all firms if, and only if, their vacancy postings have the form

$$\nu\left(e^{\gamma t}\zeta, \psi(\zeta)\right) = \overline{\theta}\psi(\zeta), \qquad \text{all } \zeta \geq z_{min,0},\ t,$$

which is the same across all firms on the one-dimensional manifold. Q.E.D.

If each firm uses the vacancy posting rate $\nu(z,l) = \overline{\theta}l$, then market tightness is indeed $\overline{\theta}$. But Lemma 3 shows only that choosing $\nu(z,l) = \overline{\theta}l$ is required for employment to remain constant at all firms on the BGP. The last step of the proof of Proposition 1 will be to show that there exists a unique value $\overline{\theta}^*$ for which, given labor market tightness $\overline{\theta}^*$, the choice $\nu(z,l) = \overline{\theta}^* l$ is *optimal* for firms.

Lemma 4. *On a BGP, if one exists, the value function* Π *for any firm on the one-dimensional manifold satisfies*

$$\begin{aligned} \Pi(z, \psi(e^{-\gamma t}z), t) =& e^{\gamma t}\Pi(e^{-\gamma t}z, \psi(e^{-\gamma t}z), 0) \\ \equiv& e^{\gamma t}\Pi_0(e^{-\gamma t}z, \psi(e^{-\gamma t}z)), \qquad \textit{all } z \geq z_{min,0},\ t. \end{aligned} \tag{1.26}$$

Proof. The optimized value of a firm is the present discounted value of its operating profits (π) less vacancy posting costs (c).

For a firm of type $(e^{\gamma t}\zeta, l) = [e^{\gamma t}\zeta, \psi(\zeta)]$, relative productivity growth is constant if and only if

$$0 = \frac{d\zeta}{dt} - \gamma\zeta = \tau\nu(e^{\gamma t}\zeta, \psi(\zeta))q(\overline{\theta}) \int_{\zeta}^{\infty} \ln\left(\frac{\zeta'}{\zeta}\right) d\Phi_0(\zeta') - \gamma, \qquad \text{all } t. \tag{1.27}$$

This condition holds for all firms on the one-dimensional manifold if, and only if, the first term on the RHS is the same across firms and equal in magnitude to the aggregate growth rate γ.

Under Assumption 1 Φ_0 has a Pareto distribution, so

$$\int_{\zeta}^{\infty} \ln\left(\frac{\zeta'}{\zeta}\right) d\Phi_0(\zeta') = \frac{1}{\chi}\left(\frac{z_{min,0}}{\zeta}\right)^{\chi}, \qquad \text{all } \zeta \geq z_{min,0}.$$

Furthermore, Lemma 3 implies that the vacancy posting rate required for constant employment is $\nu(e^{\gamma t}\zeta, \psi(\zeta)) = \overline{\theta}\psi(\zeta)$. Substituting both of these expressions into (1.27) and rearranging yields

$$\frac{\psi(\zeta)}{\zeta^{\chi}} = \frac{\gamma\chi}{\tau\overline{\theta}q(\overline{\theta})\left(z_{min,0}\right)^{\chi}}, \qquad \text{all } \zeta \geq z_{min,0}, \tag{1.28}$$

where the term on the RHS is constant across all firms on the one-dimensional manifold. By Assumption 2, the ratio on the LHS is the same for all firms on the one-dimensional manifold. Therefore, all firms on the manifold improve productivity at the same constant rate. This constant rate is γ.

Because under the assumptions of Proposition 1 all firms on the one-dimensional manifold improve productivity at the same rate and do not change firm size, then

$$\begin{aligned} \pi(e^{\gamma t}\zeta, \psi(\zeta), t) =& e^{\gamma t}\pi_0(\zeta, \psi(\zeta)), \\ c(\nu, e^{\gamma t}\zeta, \psi(\zeta), t) =& e^{\gamma t}c_0(\nu, \zeta, \psi(\zeta)), \qquad \text{all } \zeta \geq z_{min,0},\ t. \end{aligned}$$

Therefore both the profits and vacancy posting cost of each firm on the one-dimensional manifold grow at the constant rate γ. The present discounted value of future net profits for firms on the BGP is thus

$$\Pi(e^{\gamma t}z, \psi(z), t) = e^{\gamma t}\Pi_0(z, \psi(z)), \qquad \text{all } z \geq z_{min,0},\ t.$$

Q.E.D.

From (1.19),

$$r\Pi(z,l,t) = \pi(z,l,t) - l\widehat{c}(\overline{\theta})e^{\gamma t}\overline{\omega}_0 + \frac{\partial \Pi}{\partial z}\frac{dz}{dt} + \frac{\partial \Pi}{\partial l}\frac{dl}{dt} + \frac{\partial \Pi}{\partial t}, \qquad \text{all } z \geq z_{min,0},\ l \geq 0,\ t, \tag{1.29}$$

when the vacancy posting cost function takes on the form in Assumption 3, and on a BGP (if one exists) $\overline{\omega}(t) = e^{\gamma t}\overline{\omega}_0$ and $\nu(s) = \overline{\theta}\psi(\zeta)$.

When $l = \psi(e^{-\gamma t}z)$, (1.26) implies

$$\begin{aligned}\frac{\partial \Pi(z,l,t)}{\partial z} =& \frac{\partial \Pi_0(e^{-\gamma t}z,l)}{\partial z},\\ \frac{\partial \Pi(z,l,t)}{\partial t} =& e^{\gamma t}\left[\gamma \Pi_0(e^{-\gamma t}z,l) - \gamma z \frac{\partial \Pi_0(e^{-\gamma t}z,l)}{\partial z}\right], \qquad \text{all } z \geq z_{min,0},\ l \geq 0,\ t.\end{aligned}$$

On a BGP, $dz/dt = \gamma z$, so

$$\begin{aligned}\frac{\partial \Pi}{\partial z}\frac{dz}{dt} + \frac{\partial \Pi}{\partial t} =& \frac{\partial \Pi_0}{\partial z}\gamma z + e^{\gamma t}\left[\gamma \Pi_0 - \gamma z \frac{\partial \Pi_0}{\partial z}\right]\\ =& \gamma e^{\gamma t}\Pi_0, \qquad \text{all } z \geq z_{min,0},\ t.\end{aligned}$$

Also on a BGP, $dl/dt = 0$ for all z, t if $l = \psi(e^{-\gamma t}z)$.

Substituting these expressions for the derivatives of the value of a firm and dl/dt into (1.29), the value of firms on the one-dimensional manifold of the conjectured BGP satisfy

$$(\overline{r} - \gamma)\Pi_0(z,\psi(z)) = \pi_0(z,\psi(z)) - \overline{\omega}_0\psi(z)\widehat{c}(\overline{\theta}), \qquad \text{all } z \geq z_{min,0}. \tag{1.30}$$

To prove the main proposition, it remains to be shown that on the conjectured BGP the FOC for a vacancy posting in (1.21) holds for all firms on the one-dimensional manifold. In normalized form that FOC is

$$\begin{aligned}\frac{\overline{\omega}_0}{q(\overline{\theta})}\widehat{c}'(\nu(\zeta)/\psi(\zeta)) =& \frac{\partial \Pi_0(\zeta,l)}{\partial \zeta}\tau\zeta \int_{\zeta}^{\infty} \ln\left(\frac{\zeta'}{\zeta}\right) d\Phi_0(\zeta') + \frac{\partial \Pi_0(\zeta,l)}{\partial l}\int_0^{\infty} \psi(\zeta')\, d\Phi_0(\zeta')\\ =& \frac{\partial \Pi_0(\zeta,l)}{\partial \zeta}\frac{\tau}{\chi}z_{0,min}^{\chi}\zeta^{1-\chi} + \frac{\partial \Pi_0(\zeta,l)}{\partial l}, \qquad \text{all } \zeta \geq z_{min,0},\end{aligned} \tag{1.31}$$

where the derivatives of Π_0 are evaluated at $(\zeta, l) = (\zeta, \psi(\zeta))$, and where the second line uses Assumption 1 (Φ_0 is a Pareto distribution) and the fact that employment across all firms integrates to unity.

Proof. Proposition 1

Recall that since $L = 1$ and all workers search with unit intensity, average labor market tightness is, by definition,

$$\overline{\theta}(t) = \int_{\mathcal{S}} n(s,t)\, dF(s,t), \qquad \text{all } t,$$

where $n(s,t)$ is the vacancy posting rate chosen by the s-type firm. When a firm chooses its vacancy posting rate it take the tightness of the labor market, $\overline{\theta}(t)$, as given and respond optimally.

By Lemma 3, employment remains constant at every firm on the one-dimensional manifold of the BGP if, and only if, firms have the vacancy posting rate

$$n(s,t) = \overline{\theta}(t) l(s,t).$$

But the choices of the firms must also be *optimal*. Given the labor market tightness, a firm's choice must satisfy (1.31). Define

$$\lambda(\zeta; \overline{\theta}) \equiv \frac{\nu(\zeta, \psi(\zeta))}{\psi(\zeta)},$$

as the solution to (1.31) for firms on the one-dimensional manifold of the BGP, given $\overline{\theta}$. Next, it will be shown that

1. $\lambda(\zeta; \overline{\theta}) = \overline{\lambda}(\overline{\theta})$ is the same for all firms (i.e. all $\zeta \geq z_{min,0}$), and
2. there exists a unique value $\overline{\theta}^*$ with the property that $\overline{\lambda}(\overline{\theta}^*) = \overline{\theta}^*$.

To show that all firms follow the same vacancy posting strategy it is necessary to develop expressions for $\partial \Pi_0 / \partial \zeta$ and $\partial \Pi_0 / \partial l$. For the former, differentiate (1.30) with respect to ζ to get

$$(\overline{r}-\gamma)\left[\frac{\partial \Pi_0(\zeta,\psi(\zeta))}{\partial \zeta}+\frac{\partial \Pi_0(\zeta,\psi(\zeta))}{\partial l}\psi'(\zeta)\right]=\frac{\partial \pi_0(\zeta,\psi(\zeta))}{\partial \zeta}+\frac{\partial \pi_0(\zeta,\psi(\zeta))}{\partial l}\psi'(\zeta)$$
$$-\overline{\omega}_0\widehat{c}(\overline{\theta})\psi'(\zeta),\qquad \text{all } \zeta\geq z_{min,0},$$

or

$$\frac{\partial \Pi_0}{\partial \zeta}=\frac{1}{\overline{r}-\gamma}\left[\frac{\partial \pi_0}{\partial \zeta}+\left(\frac{\partial \pi_0}{\partial l}-\overline{\omega}_0\widehat{c}(\overline{\theta})\right)\psi'(\zeta)\right]-\frac{\partial \Pi_0}{\partial l}\psi'(\zeta),\qquad \text{all } \zeta\geq z_{min,0}, \tag{1.32}$$

where the derivatives of π_0 are evaluated at $(\zeta,l)=(\zeta,\psi(\zeta))$.

Substitute (1.32) into (1.31) to write the vacancy posting FOC as

$$\begin{aligned}\frac{\overline{\omega}_0}{q(\overline{\theta})}\widehat{c}'(\lambda(\zeta;\overline{\theta}))=&\frac{1}{\overline{r}-\gamma}\left[\frac{\partial \pi_0}{\partial \zeta}+\left(\frac{\partial \pi_0}{\partial l}-\overline{\omega}_0\widehat{c}(\overline{\theta})\right)\psi'(\zeta)\right]\frac{\tau}{\chi}\zeta_{0,min}^{\chi}\zeta^{1-\chi}\\&+\frac{\partial \Pi_0(\zeta,l)}{\partial l}\left(1-\psi'(\zeta)\frac{\tau}{\chi}\zeta_{0,min}^{\chi}\zeta^{1-\chi}\right)\\=&\frac{\tau z_{0,min}^{\chi}}{\overline{r}-\gamma}\left[\frac{\partial \pi_0}{\partial \zeta}\frac{\zeta^{1-\chi}}{\chi}+\left(\frac{\partial \pi_0}{\partial l}-\overline{\omega}_0\widehat{c}(\overline{\theta})\right)\psi_0\right]\\&+\frac{\partial \Pi_0(\zeta,l)}{\partial l}\left(1-\tau\psi_0 z_{0,min}^{\chi}\right),\qquad \text{all } \zeta\geq z_{min,0},\end{aligned} \tag{1.33}$$

where the second part uses Assumption 2 which implies $\psi'(\zeta)\zeta^{1-\chi}/\chi=\psi_0$, and all derivative terms are evaluated at $(\zeta,l)=(\zeta,\psi(\zeta))$.

Because (1.30) only holds for firms on the one-dimension manifold, it cannot be used to derive an expression for $\partial\Pi_0/\partial l$. Instead, a direct approach is required.

Let $\Delta_{l0}>0$ be a small perturbation in firm size at date $t=0$. Suppose that a firm with initial productivity z begins with $\psi(z)+\Delta_{l0}$ workers. Also suppose that at each date the firm pays the same wage and posts the same number of vacancies as a firm with initial productivity z and $\psi(z)$ workers. Then the firm with the incremental amount of additional workers has the same hiring rate, and the stock of knowledge evolve over time as if it was the firm with $\psi(z)$ workers on the one-dimension manifold. Over time, the incremental workers are poached by other firms at the rate $\overline{\theta}q(\overline{\theta})$. Thus the increment of extra workers remaining at date t is $\Delta(t) = e^{\overline{\theta}q(\overline{\theta})t}\Delta_{l0}$.

At date t, the firm's incremental profit flow per worker is

$$\left.\frac{\partial\pi(\zeta,l,t)}{\partial l}\right|_{l=\psi(e^{-\gamma t}\zeta)} = e^{\gamma t}\left.\frac{\partial\pi_0(e^{-\gamma t}\zeta,l)}{\partial l}\right|_{l=\psi(e^{-\gamma t}\zeta)}, \qquad \text{all } \zeta \geq z_{min,0},\ t.$$

Hence at the interest rate $\overline{r}$, the present discounted value of the incremental profit flows resulting from the increment of workers is

$$\frac{\partial\Pi_0}{\partial l} = \frac{1}{\overline{r}-(\gamma-\overline{\theta}q(\overline{\theta}))}\frac{\partial\pi_0}{\partial l}, \qquad \text{all } \zeta \geq z_{min,0}, \tag{1.34}$$

where both derivative terms are evaluated at $(\zeta,l) = (\zeta,\psi(\zeta))$.

Recall from (1.19) that hiring has been optimized by firms. Hence the envelop theorem ensures that, to a first-order approximation, the term in braces in (1.19) does not change in response to the small perturbation Δ_{l0}.

Substitute (1.34) into (1.33) to write the FOC as

$$\begin{aligned}\frac{\overline{\omega}_0}{q(\overline{\theta})}\widehat{c}'(\lambda(\zeta;\overline{\theta})) =& \frac{\tau z_{min,0}^{\chi}}{\overline{r}-\gamma}\left[\frac{\partial\pi_0}{\partial\zeta}\frac{\zeta^{1-\chi}}{\chi} + \left(\frac{\partial\pi_0}{\partial l} - \overline{\omega}_0\widehat{c}(\lambda(\zeta;\overline{\theta}))\right)\psi_0\right] \\ &+ \frac{1-b_0}{\overline{r}-(\gamma-\overline{\theta}q(\overline{\theta}))}\frac{\partial\pi_0}{\partial l} \\ =& \frac{\tau z_{min,0}^{\chi}}{\overline{r}-\gamma}\left[\frac{\partial\pi_0}{\partial\zeta}\frac{\zeta^{1-\chi}}{\chi} - \overline{\omega}_0\widehat{c}(\lambda(\zeta;\overline{\theta}))\psi_0\right] \\ &+ \frac{\partial\pi_0}{\partial l}\left(\frac{b_0}{\overline{r}-\gamma} + \frac{1-b_0}{\overline{r}-(\gamma-\overline{\theta}q(\overline{\theta}))}\right), \qquad \text{all } \zeta \geq z_{min,0},\end{aligned} \tag{1.35}$$

where $b_0 \equiv \tau\psi_0 z_{min,0}^{\chi}$ is a constant.

To show the same value of $\lambda(\zeta;\overline{\theta})$ satisfies the FOC (1.35) for all firms on the one-dimensional manifold, it is suffices to show that

$$\frac{\partial \pi_0}{\partial \zeta}\zeta^{1-\chi} \qquad \text{and} \qquad \frac{\partial \pi_0}{\partial l}$$

are the same for all $(\zeta, \psi(\zeta))$.

For notational convenience, define $\eta \equiv 1 - 1/\rho$. Recall that

$$\begin{aligned}
\pi(z,l,t) =& (1-\widehat{\beta})R(z,l,t) \\
=& (1-\widehat{\beta})P(t)Y(t)^{1/\rho}(zl^{\alpha})^{\eta}, \qquad \text{all } z \geq z_{min,0},\ l \geq 0,\ t.
\end{aligned}$$

Thus on a BGP of the type conjectured,

$$\begin{aligned}
\pi(z,l,t) =& e^{\gamma t}\pi_0(e^{-\gamma t}z, l) \\
=& e^{\gamma t}x_0[e^{-\gamma t}zl^{\alpha}]^{\eta}, \qquad \text{all } z \geq z_{min,0},\ l \geq 0,\ t,
\end{aligned}$$

where x_0 is a constant. When evaluated at $t = 0$,

$$\pi_0(\zeta, l) = x_0\zeta^{\eta}l^{\alpha\eta}, \qquad \text{all } \zeta \geq z_{min,0},\ l \geq 0. \tag{1.36}$$

Taking the derivative of (1.36) with respect to ζ and then evaluating at $(\zeta, \psi(\zeta))$ yields

$$\begin{aligned}
\left.\frac{\partial \pi_0(\zeta,l)}{\partial \zeta}\right|_{l=\psi(\zeta)} =& x_0\eta\zeta^{\eta-1}\psi(\zeta)^{\alpha\eta} \\
=& x_0\eta\psi_0^{\alpha\eta}\zeta^{\eta-1}\zeta^{\chi\alpha\eta}, \qquad \text{all } \zeta \geq z_{min,0}.
\end{aligned}$$

Recall that $\chi = \eta/(1-\alpha\eta)$. Hence,

$$\begin{aligned}
\left.\frac{\partial \pi_0(\zeta,l)}{\partial \zeta}\right|_{l=\psi(\zeta)} \zeta^{1-\chi} =& x_0\eta\psi_0^{\alpha\eta}\zeta^{\eta+\chi(\alpha\eta-1)} \\
=& x_0\eta\psi_0^{\alpha\eta}, \qquad \text{all } \zeta \geq z_{min,0}.
\end{aligned} \tag{1.37}$$

Therefore, $(\partial\pi_0/\partial l)\zeta^{1-\chi}$ is constant across all firms on the one-dimensional manifold.

Similarly, taking the derivative of (1.36) with respect to l and then evaluated at $(\zeta, \psi(\zeta))$ yields

$$\begin{aligned}
\left.\frac{\partial\pi_0(\zeta, l)}{\partial l}\right|_{l=\psi(\zeta)} =& x_0\zeta^\eta \alpha\eta\psi(\zeta)^{\alpha\eta-1} \\
=& \alpha\eta x_0\psi_0^{\alpha\eta-1}, \qquad \text{all } \zeta \geq z_{min,0}.
\end{aligned} \tag{1.38}$$

So $\partial\pi_0/\partial l$ is also constant across firms on the one-dimensional manifold.

Substituting (1.37) and (1.38) into (1.35), the FOC can be rewritten as

$$\begin{aligned}
\frac{\overline{\omega}_0}{q(\overline{\theta})}\widehat{c}'(\lambda(\zeta;\overline{\theta})) =& \frac{b_0}{\overline{r}-\gamma}\left[\frac{x_0\eta\psi_0^{\alpha\eta-1}}{\chi} - \overline{\omega}_0\widehat{c}(\lambda(\zeta;\overline{\theta}))\right] \\
&+ \alpha\eta x_0\psi_0^{\alpha\eta-1}\left(\frac{b_0}{\overline{r}-\gamma} + \frac{1-b_0}{\overline{r}-(\gamma-\overline{\theta}q(\overline{\theta}))}\right), \qquad \zeta \geq z_{min,0},
\end{aligned}$$

or

$$\begin{aligned}
\frac{\overline{\omega}_0}{q(\overline{\theta})}\widehat{c}'(\lambda(\zeta;\overline{\theta})) + \frac{b_0}{\overline{r}-\gamma}\overline{\omega}_0\widehat{c}(\lambda(\zeta;\overline{\theta})) =& \frac{b_0}{\overline{r}-\gamma}\frac{x_0\eta\psi_0^{\alpha\eta-1}}{\chi} \\
&+ \alpha\eta x_0\psi_0^{\alpha\eta-1}\left(\frac{b_0}{\overline{r}-\gamma} + \frac{1-b_0}{\overline{r}-(\gamma-\overline{\theta}q(\overline{\theta}))}\right), \zeta \geq z_{min,0}.
\end{aligned} \tag{1.39}$$

Note that the value on the RHS of (1.39) is positive and independent of ζ. By Assumption 3, the LHS is strictly increasing in λ, takes the value zero at $\lambda = 0$, and diverges as $\lambda \to \infty$. Therefore, for any fixed $\overline{\theta}$, there is a unique value of λ that satisfies (1.39). Furthermore, this value of λ is independent of ζ. Hence all firms will choose vacancy postings to satisfy

$$\lambda(\zeta;\overline{\theta}) = \overline{\lambda}(\overline{\theta}), \qquad \text{all } \zeta \geq z_{min,0}.$$

The second step in the proof is to show that there exists a unique value of $\overline{\theta}^*$ for which

$\overline{\lambda}(\overline{\theta}^*) = \overline{\theta}^*$.

From Lemma 3, when all firms choose the same value of λ then $\overline{\theta} = \overline{\lambda}(\overline{\theta})$. The FOC for all firms on the BGP then becomes

$$\frac{\overline{\omega}_0}{q(\overline{\theta})}\widehat{c}'(\overline{\theta}) + \frac{b_0}{\overline{r}-\gamma}\overline{\omega}_0\widehat{c}(\overline{\theta}) = \frac{b_0}{\overline{r}-\gamma}\frac{x_0\eta\psi_0^{\alpha\eta-1}}{\chi} + \alpha\eta x_0\psi_0^{\alpha\eta-1}\left[\frac{b_0}{\overline{r}-\gamma} + \frac{1-b_0}{\overline{r}-(\gamma-\overline{\theta}q(\overline{\theta}))}\right]. \tag{1.40}$$

Define two functions $Z_L(\overline{\theta})$ and $Z_R(\overline{\theta})$ by the LHS and RHS of (1.40) respectively.

It will be shown that

$$\begin{aligned} &Z_L(0) = 0, && Z_L'(\overline{\theta}) > 0, && \text{and} \lim_{\overline{\theta}\to\infty} Z_L(\overline{\theta}) = \infty, \\ &Z_R(0) = \overline{Z}_R, && Z_R'(\overline{\theta}) < 0, && \text{and} \lim_{\overline{\theta}\to\infty} Z_R(\overline{\theta}) = \underline{Z}_R, \end{aligned}$$

where $0 < \underline{Z}_R < \overline{Z}_R$ are finite. It then follows immediately that there exists a unique value of $\overline{\theta}^*$ such that $\overline{\lambda}(\overline{\theta}^*) = \overline{\theta}^*$.

First consider $Z_L(\overline{\theta})$. The function $q(\overline{\theta}) = \overline{q}\overline{\theta}^{-\mu}$ is strictly decreasing in $\overline{\theta}$ and diverges as $\overline{\theta} \to 0^+$. By Assumption 3, $\widehat{c}(0) = \widehat{c}'(0) = 0$. Hence $Z_L(0) = 0$, and $Z_L(\overline{\theta})$ diverges as $\overline{\theta} \to \infty$. The derivative of Z_L is given by

$$Z_L'(\overline{\theta}) = \widehat{c}'(\overline{\theta})\overline{\omega}_0\left(\frac{\overline{\theta}^{\mu}}{\overline{q}}\left[\frac{\widehat{c}''(\overline{\theta})}{\widehat{c}'(\overline{\theta})} + \mu\overline{\theta}\right] + \frac{b_0}{\overline{r}-\gamma}\right) > 0, \qquad \text{all } \overline{\theta} \geq 0.$$

By Assumption 3, the term in the braces is positive implying that the value of $Z_L(\overline{\theta})$ is monotonically increasing for all $\overline{\theta} \geq 0$.

For the function $Z_R(\overline{\theta})$, since $\mu \in (0,1)$, the term $\overline{\theta}q(\overline{\theta}) = \overline{q}\overline{\theta}^{1-\mu}$ takes the value zero at $\overline{\theta} = 0$. At $\overline{\theta} = 0$ the value of the Z_R is

$$\overline{Z}_R \equiv Z_R(0) = \frac{x_0\eta\psi_0^{\alpha\eta-1}}{\overline{r}-\gamma}\left(\frac{b_0}{\chi} + \alpha\right) > 0.$$

As $\overline{\theta}$ diverges, the the last term in the brackets of value of (1.40) vanishes, so Z_R converges to the finite value

$$\underline{Z}_R \equiv \lim_{\overline{\theta}\to\infty} Z_R(\overline{\theta}) = \frac{x_0 \eta \psi_0^{\alpha\eta-1}}{\overline{r}-\gamma}\left(\frac{b_0}{\chi}+\alpha b_0\right) < \overline{Z}_R.$$

Finally, the derivative of $Z_R(\overline{\theta})$ is

$$Z_R'(\overline{\theta}) = -X_0 \frac{(1-\mu)\overline{q}\overline{\theta}^{-\mu}}{\left[r-\gamma+\overline{q}\overline{\theta}^{1-\mu}\right]^2} < 0 \qquad \text{all } \overline{\theta} \geq 0,$$

where X_0 is a constant. Q.E.D.

Let the unique value of $\overline{\theta}$ that satisfies (1.40) be denoted by $\overline{\theta}^*$. The proof of the proposition implies that when the initial distribution of firms is one-dimensional and satisfies the assumptions made in Proposition 1, all firms in the competitive equilibrium follow the optimal vacancy posting rule

$$\nu(s,t) = \overline{\theta}^* \psi(e^{-\gamma t} z), \qquad \text{all } s \in \mathcal{S}, t.$$

In such an environment, each firm type maintains its initial firm size, and the productivity level of all firms grows at the constant rate γ, generating a BGP.

1.6 Conclusions

The endogenous growth model developed in this paper uses learning from new hires to provide a channel for the spillover of knowledge between firms. Workers passively absorb knowledge while working and are able to take some of this knowledge to their new employer. Firms optimally choose their vacancy posting rate, which determined their exposure to new knowledge. This structure endogenizes the learning process that is usually treated as exogenous in traditional endogenous growth models with knowledge spillover.

The main result of this paper shows that there is a set of assumptions regarding the initial distribution of firm types and the vacancy posting cost function that are sufficient for the competitive equilibrium to be a balanced growth path. Along this balanced growth path aggregate output grows at a constant rate and the size/employment of each firm is constant. Therefore, while knowledge spillovers between firms are likely to occur through many different channels, the diffusion of knowledge through new hires is capable of explaining long-run sustainable growth and could be used as a structural explanation of the exogenous learning process in standard models of endogenous growth.

CHAPTER 2

FIRM PRODUCTIVITY GROWTH AND ITS RELATIONSHIP TO THE KNOWLEDGE OF NEW WORKERS

Michael Kirker
University of Chicago

Lynda Sanderson
Ministry of Business, Innovation and Employment

2.1 Abstract

Using linked employer-employee data from New Zealand, a firm's productivity growth is related to the firm's exposure to outside knowledge through the hiring of new workers with previous experience at other firms. The estimated relationship between productivity growth and hiring patterns is compared to the predictions implied by both the worker quality and knowledge spillover channels. While not a causal relationship, the multi-factor productivity results are consistent with the productivity of a worker's previous employer acting as a signal of the unmeasured quality/inherent-knowledge of the new worker. When firm productivity is measured in terms of labor productivity, the results are also consistent with new workers spilling over knowledge that allows the hiring firm to adopt more capital intensive production techniques when the new worker was previously employed by a more productive firm.

2.2 Disclaimer

The results in this paper are not official statistics, they have been created for research purposes from the Integrated Data Infrastructure (IDI), managed by Statistics New Zealand.

The opinions, findings, recommendations, and conclusions expressed in this paper are those of the author(s), not Statistics NZ, the Treasury, or the Ministry of Business, Innovation and Employment.

Access to the anonymised data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this paper have been confidentialised to protect these groups from identification.

Careful consideration has been given to the privacy, security, and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the Privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz.

The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information may be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes.

Any person who has had access to the unit record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to secrecy. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

2.3 Introduction

Improvements in innovation and productivity have long been regards by the economic literature as the key drive of long-run economic growth. According to New Zealand firms, new staff are an important source of ideas for new innovation. Every two years, Statistics New Zealand surveys businesses about their innovation practices in the Innovation Module of the Business Operations Survey (BOS), a nationally representative survey of firms. In 2013, 46 percent of responding businesses reported that they had implemented some form of innovation in the prior two years.[1] Of those businesses that reported carrying out some form of innovation, 52 percent reported that new staff were an important source of ideas for the innovation that was carried out.

Despite firms report that new staff are an important source of ideas for innovation, the mechanism by which this benefit occurs is not well understood. The economic literature has proposed many theoretical channels through which the knowledge of new workers can influence the hiring firm's productivity such as knowledge spillover, work-firm match quality, and worker skill. However, there is scarce empirical support for which, if any, of these channels actually exist in the data.

This paper aims to improve our understanding of the channels through which firms benefit from the knowledge and ideas brought to the firm by new hires. It studies in detail the relationship between characteristics of new hires and measurable productivity growth at the hiring firm. The analysis uses panel data that matches the full population of businesses to their workers to examine how growth in a firm's productivity is related to both the skill/quality of new workers and the knowledge new workers may have acquired working at their previous firms. These empirical relationships are then compared to the predictions made by two theoretical channels that the literature has used to relate firm productivity and labor mobility. Namely, a productive knowledge spillover channel and an unmeasured worker quality channel.

The analysis expands on the previous empirical literature in two key areas. First, due to restrictions in data availability, the previous literature has predominantly focused on examining knowledge spillovers in manufacturing industries, where revenue and cost data

1. The types of innovation asked about were: (i) product innovation: "did this business introduce onto the market any new or significantly improved goods or services?"; (ii) process innovation: "did this business implement any new or significantly improved operational processes (ie methods of producing or distributing goods or services)?"; (iii) organisational innovation: "did this business implement any new or significantly improved organisational/managerial processes (ie significant changes in this businesses strategies, structures or routines)?"; and (iv) marketing innovation: "did this business implement any new or significantly improved sales or marketing methods which were intended to increase the appeal of goods or services for specific market segments or to gain entry to new markets?"

are more readily available. The data used in this paper provides coverage of firms in all industries of the measurable economy.[2] Second, another limitation common in the previous literature is only observing employment data at a particular date each year. The employment information available in the New Zealand data is observed at the monthly frequency. Not only does this frequency of observation allow for more precision in the timing of job starts/finishes, but it also ensures that we are able to capture all jobs that a worker undertakes (and not just the jobs where the worker was employed at a particular date each year).

In addition to the empirical contributions, this paper builds upon the theoretical base of the previous literature by adding additional controls to the model that have not generally been used. Previous papers have tended to only control for the hiring intensity of workers from sources for which it is possible to measure productivity (e.g. only hires from other manufacturing firms). A firm's decision to hire workers from sources within the scope of productivity analysis are likely to be correlated with its decision to hire from sources outside that scope. Without controlling for this effect, the estimated size of the productive knowledge spillover effect may be biased. Due to the richness of the New Zealand data, it is possible to introduce controls for hires outside of the scope of our analysis (such as hires from non-market firms, or new entrants to the labor market) as an attempt to control for the possibility of knowledge spillovers from these other sources.

To help distinguish between worker quality and knowledge spillover effects, the analysis makes a distinction between the intensive margin of hiring, proxied by the productivity of the new worker's previous firm, and the extensive margin of hiring, how many workers were hired. Overall, the results from the regressions show that when a firm hires new workers, the productivity of the workers' previous employer is significantly correlated with the productivity gains at the hiring firm following the new hires, even after controlling for changes in the (measured) quality of the firm's labor force.

2. The measured sector of the economy is defined by Statistics New Zealand as industries that mainly contain enterprises that are market producers.

When firm productivity is measured in terms of multi-factor productivity, higher productivity growth in the hiring firm is positively correlated with a higher the average productivity level of the private-for-profit firms that new workers are sourced from. The size of the expected increase in productivity growth is the same irrespective of whether the firm increases the average productivity of the less productive firms it hires from or increases the average productivity of the more productive firms it hires from. However, when firm-level productivity is measured in terms of labor productivity (value-added per worker), raising the average productivity of the more productive firms that workers are hired from is associated with a larger expected productivity gain at the hiring firm than raising the average productivity of the less productive firms that workers are hired from. In addition, when the flow of new workers into the hiring firm is further sub-divided based on worker and firm characteristics, the variation in productivity gains and losses from these various sub-divisions is larger when using value-added as the productivity measures than when using multi-factor productivity measures.

The overall pattern of correlations described above does not point to a unique channel through which firms benefit from new workers. Instead, the results are consistent with a story in which new workers affect productivity in the hiring firm through both knowledge spillovers and changing the unmeasured worker quality within the firm. The results also suggest that if productive knowledge spillover is one of the causal drivers of the firm's productivity growth, then the type of knowledge that spills over between firms relates to technology knowledge, which allows firms to take advantage of more capital-intensive production techniques, rather than multi-/total-factor productivity knowledge which would allow for the more efficient utilisation of existing inputs. While it is not possible to definitively conclude the direction of causality in these relationships, these finding do appear to be robust to the attempts we can make to control for causality, suggesting that at least some part of the relationship between the knowledge of new workers and the subsequent productivity growth at the hiring firm is likely to run in the direction from workers to firm productivity.

The remainder of the paper is structured as follows. Section 2.4 discusses how this paper fits into the existing literature. Section 2.5 discusses the model used for the analysis. Section 2.6 details the data sources used. Section 2.7 presents the results of the analysis. Section 2.8 concludes.

2.4 Related Literature

Typically in the empirical literature, the measure of a firm's exposure to new productive knowledge from workers is proxied by the share of new workers at the hiring firm. Using Danish data on several industries, Parrotta and Pozzoli (2012) find that the number of hires of new highly-educated workers — who are likely to be carriers of knowledge between firms — is correlated with productivity growth in the hiring firm. Similarly, Serafinelli (2015) shows that the productivity of Italian manufacturing firms improves when hiring workers from high wage premium firms (a proxy for high productivity firms).

The analytical approach taken in this paper most closely relates to that used by Stoyanov and Zubanov (2012) who use the notion of a 'productivity gap' — the difference between the hiring firm's productivity and the productivity of the new workers' previous employers — as a measure of the hiring firm's exposure to new knowledge. Their analysis shows that for Danish manufacturing firms, hiring new workers from more productive firms benefits the hiring firm's productivity, while hiring new workers from less productive firms does not have a significant effect on the hiring firm's productivity. These correlations match the predictions of the knowledge spillover channel.

Empirically, the analysis in this paper extends that of Stoyanov and Zubanov (2012) by using data that covers all industries in the measured economy. In addition, the employment data used is able to capture all job spells, not just those observed at a particular date each year. Theoretically, the model used in this paper builds on that of Stoyanov and Zubanov (2012) by including controls for hires from various sources outside the scope of the productivity analysis, and relates the productivity gap to growth in the hiring firm's productivity, rather than the level of productivity. We believe such an approach provides a better fit with the way multi-factor productivity is typically computed in the data.

Not all papers in the literature find support for labor mobility being a channel for productive knowledge spillover. Stockinger and Wolf (2016) find that for multiple German industries the number of new workers hired from superior (defined as higher-paying) establishments does not have a significant effect on the hiring establishment's productivity. However, hiring more workers from lower-paying establishments is associated with productivity gain. Their findings suggest the productivity gains associated with new hires are more consistent with an assortative matching process — where higher (lower) skilled workers move up (down) the firm productivity ladder over time.

Motivated by this finding, we expand the scope of our analysis to also consider other possible channels beyond knowledge spillover through which new workers will benefit the hiring firm. The empirical correlations are then compared to these various channels as a way to help choose between the competing stories for how firms benefit from the knowledge of new workers in New Zealand.

In the context of the New Zealand literature, the relationship between firm-level innovation and the characteristics of new workers has primarily focused on flows of migrants. McLeod et al. (2014) find that a higher proportion of recent migrants within the firms' workforce is correlated with a higher probability of self-reporting innovation in the BOS. Sin et al. (2014) find that hiring high skilled foreigners raises the probability that a firm will self-report both exporting and innovating. While these findings only represent correlations, they are consistent with the causal story of foreign knowledge spillover through the international migration of labor.

The analysis carried out in this paper expands upon this previous New Zealand literature by considering all labor flows, not just those related to international migration. Also, firm-level innovation is views through the lens of measurable productivity rather than the binary self-reported BOS responses. This provides an indication of the magnitude of innovation that is occurring within businesses.

2.5 The Model

The analytical framework consists of two stages. The first stage is to derive firm-specific measures of productivity. The second stage is to explicitly model the relationship between growth in the hiring firm's productivity and the firm's exposure to productive knowledge and skills brought to the firm by new workers.

2.5.1 Measuring Productivity

This paper considers a range of alternate measures of firm productivity.[3] The first measure considered is labor productivity, which has the advantages of being straightforward to compute, and allows for direct comparisons on a like-for-like basis between firms in different industries and firms that employ different levels of inputs. Labor productivity is calculated as the real value-added (value of the final output less materials) per full-time equivalent (FTE) worker. More formally, let $A_{i,t}$ denote labor productivity for firm i in year t. Labor productivity is then defined as

$$A_{i,t} = \frac{Y_{i,t} - M_{i,t}}{L_{i,t}}, \tag{2.1}$$

where $Y_{i,t}$ denotes the real value of the firm's output in year t, $M_{i,t}$ denotes the real value of material inputs into the production process, and $L_{i,t}$ is the measure of labor input in FTE units.

Also considered are various measures of multi-factor productivity (MFP). Multi-factor productivity controls for changes in other factor inputs (such as capital) and returns to scale. However, the measure of MFP is dependent upon the functional form of the benchmark

3. In addition to firm productivity, some of the analysis examines the capital-labor ratio as a measure of input intensity.

production function that is chosen. Formally, let $Y_{i,j,t}$ denote the output of firm i, in industry j, at time t. The firm's output can be expressed as

$$Y_{i,j,t}(L, K, M) = A_{i,t} F_{j,t}(L, K, M), \tag{2.2}$$

where $A_{i,t}$ is the firm's MFP, $F_{j,t}(\cdot)$ is the production function technology used by industry j at date t, and L, K, and M, are the firm's choice of labor, capital, and materials respectively.[4] Given information on the firm's level of output, inputs, and a functional form for the production function (e.g. Cobb-Douglas technology), (2.2) can be used to estimate the level of MFP for the firm ($A_{i,t}$) as a residual. One limitation of this approach is that MFP is a relative measure, and like-for-like comparisons can only be made between firms using the same production function benchmark.

2.5.2 Modeling Productivity Growth

This paper is primarily focused on how the knowledge new workers may have influences the productivity at a firm. The analysis to follow makes a distinction between two types of knowledge. The first type is knowledge that is intrinsic to the structure/operation of a firm such as the managerial, marketing, or production methods employed. Such knowledge is fairly invariant to the specific workers employed by the firm at any point in time. If one worker leaves, a new worker can be hired and placed in the vacancy role and the processes used by the firm will be unaffected. What is more, such knowledge is non-rival and can be copied by other firms.

The second type of knowledge is knowledge that is intrinsic to the skill or quality of workers such as education or innate worker ability. This type of knowledge can only be used by the firm when the worker is present. If the worker leaves the firm, they take this type of

4. Throughout the rest of this paper, $A_{i,t}$ and the term 'firm productivity' will be used to refer to the firm's productivity measured either as labor productivity or MFP.

knowledge will them, thereby lowering productivity at the firm.[5]

Treating the firm's productivity as the Solow residual, the change in the firm's productivity can be related to the change in the knowledge employed by the firm using the relationship

$$\Delta \ln A_{i,j,t} = \Delta I^{\text{firm}}_{i,t} + \gamma \Delta Q_{i,t} + \eta_{i,t}, \tag{2.3}$$

where $\Delta I^{\text{firm}}_{i,t}$ is the change in the stock of intrinsic knowledge of the firm, $\Delta Q_{i,t}$ is the change in average quality/knowledge of the workers, and $\eta_{i,t}$ is a residual capturing the change in all other productivity factors.

Some component of worker quality/knowledge may be unobservable to the econometrician, and hence unmeasured worker quality could affect $\Delta \ln A_{i,j,t}$ though factors other than $\Delta Q_{i,t}$ if it is orthogonal to observed worker quality. This issue will be discussed later.

By assumption, the stock of intrinsic firm knowledge improves as the firm receives exposure to new productive knowledge brought to the firm by new workers. Therefore we model the change in firm intrinsic knowledge by a proxy for the firms exposure to new ideas, $\Delta I^{\text{firm}}_{i,t} = \text{Exposure}_{i,t}$ that will be defined and discussed shortly. $\Delta Q_{i,t}$ will fluctuate with the observed quality of the average worker at each firm. The regression analysis also augmented (2.3) with a series of other control variables for factors that may also influence a firm's productivity. The resulting equation that will be used in the regression analysis is given by the following first-difference representation of a dynamic panel model

$$\begin{aligned} \Delta \ln A_{i,j,t} = {} & \text{Exposure}_{i,t} + \gamma \Delta Q_{i,t} + \delta \Delta \text{ExTurn}_{i,t} + \sum_{l=1}^{L} \alpha_{A,l} \Delta \ln A_{i,j,t-l} \\ & + \theta_{j,t} + \varepsilon_{i,t}, \end{aligned} \tag{2.4}$$

where $\text{ExTurn}_{i,t}$ is a measure of the excess turnover in the firm, $\sum_{l=1}^{L} \beta_{A,l} \Delta \ln A_{i,j,t-l}$ is

5. The production function generally only controls for the quantity of labor, not the quality.

a series of lagged autoregressive terms, $\theta_{j,t}$ is an industry-year fixed effect, and $\varepsilon_{i,t}$ is the regression residual term.

Excess labor turnover – a measure of the number of worker accessions and separations over and above those required to give effect to the firm's net change in employment – is included as a control because labor turnover can be disruptive to a firm when a significant amount of resources are needed to replace/train workers. Hence high labor turnover may be correlated with low levels of productivity and output.[6]

Lags of past productivity changes are included as the firm's past productivity performance can also affect productivity through influencing the investment and hiring decisions made both today and in the past. Since we are not able to explicitly model all other potential sources of new knowledge, other factors that influence firm productivity are implicitly assumed to be time-invariant (and captured by a firm-specific fixed effect), or random and independent of the other regressors (and captured by the random error term). Finally, the industry-year fixed effect soaks up any industry-wide trends in firm productivity that may remain in the data.

Dynamic panel models are known to suffer from Nickell (1981) bias that creates a correlation between the lagged productivity term and the regression's residual. While first differencing does not directly address the Nickell bias ($\Delta \ln A_{i,j,t-1}$ is still correlated with $\varepsilon_{i,t} = \Delta v_{i,t}$), it does allow us to use $\ln A_{i,t-2}$ is a natural instrument for $\Delta \ln A_{i,j,t-1}$ to control for some of the bias. More sophisticated approaches such as Blundell and Bond (1998) and other adaptations of the Arellano-Bond estimator are also suitable for the estimate of the model described above.

One limitation of the modelling approach adopted here is that it does not allow for the systematic depreciation of productive knowledge at different rates across firms. In reality certain knowledge is likely to become obsolete over time. However, modelling the deprecia-

6. Results do not differ significantly if the share of workers who exit the firm is used in place of excess turnover as a proxy for the disruption of labor turnover.

tion of productive knowledge within the firm is challenging and would require many strong assumptions to be made. As a result, we instead rely on the auto-regressive terms and idiosyncratic shocks to proxy for this process.

2.5.3 Exposure to Outside Knowledge

The firm's exposure to outside knowledge is assumed to take place through the hiring of new workers with experience at other firms. This is affected by both an intensive margin (the quality of knowledge) and an extensive margin (how many new workers). For reasons that are discussed later, all new productive knowledge is assumed to take one period (a year) to be implemented in the hiring firm before it affects the firm's productivity. Therefore, it is the workers hired in period $t-1$ that affect productivity in period t through the exposure to outside knowledge.

The baseline specification used to model the hiring firm's exposure to outside knowledge is given by

$$\begin{aligned} \text{Exposure}_{i,t} &= \beta_{agg} \frac{\sum_{n \in \mathcal{N}_{i,t-1}} [\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{H_{i,t-1}} \frac{H_{i,t-1}}{L_{i,t-1}} \\ &\quad + \sum_{s \in \mathcal{S}_{i,t-1}} \lambda_s \frac{H_{i,s,t-1}}{L_{i,t-1}}. \end{aligned} \tag{2.5}$$

The first term in the right-had side is similar to what Stoyanov and Zubanov (2012) refer to as the 'productivity gap'. $\mathcal{N}_{i,t-1}$ represents the set of all new hires by firm i at time $t-1$ from firms for which we are able to measure productivity. With a slight abuse of notation, let new hire n's previous employer also be denoted as firm n.

For each worker n who joins firm i during $t-1$, the date they departed their previous employer is denoted as date $\tau(n) \leq t-1$. The hiring firm's (i's) exposure to new knowledge from worker n depends upon the difference between the productivity between the worker's previous employer at the time they left, $\ln(A_{n,\tau(n)})$, and the hiring firm's productivity at the time worker n joins the firm $\ln(A_{i,t-1})$.

The new knowledge from each new worker is averaged over all hires with observed productivity gaps and then multiplied by the share of these workers employed by the firm (H/L).

In the data it is not possible to measure the productivity of the previous employer for every new worker. For example, some new hires may be new entrants to the labor market (and hence have no previous employer), or may come from firms for which productivity cannot be measured in the data (e.g. non-profit firms). As a result, the productivity gap that the econometrician can observe represents only a fraction of the potential exposure to outside knowledge coming from all workers. It is important to control for the entry of workers from these sources because a firm's decisions regarding hiring from firms for which the productivity gap can be computed is likely to be correlated with their decisions to hire new workers from sources for which it cannot. Therefore, failing to control for hires from these other sources would bias our estimate of the marginal effect of new productivity knowledge to the hiring firm, β_{agg}.

Despite not being able to measure the productivity of all firms in the economy, it is possible to identify in the data the reason why the productivity of the worker's previous employer is unavailable. In the second term on the right-hand-side of (2.5) let $\mathcal{S}_{i,t-1}$ denote the set of sources from which the hiring firm obtains its new workers. $H_{i,s,t-1}/L_{i,t-1}$ represents the number of hires from source $s \in \mathcal{S}_{i,t-1}$ as a fraction of the hiring firm's labor force size (the hiring intensity from source s). This term will be refered to as the hiring intensity. For new hires from sources for which it is not possible to measure the productivity gap, the parameter λ_s represents the average knowledge spillover from source s in terms of the productivity change at the hiring firm.

For hires from sources for which it is possible to measure the productivity gap, the separate productivity gap and hiring intensity terms allow for the distinction between the effects of the intensive (productivity gap) and extensive (hiring intensity) margins of knowledge exposure. In this sense, we expect that on average, the more productive the source firms that a firm is hiring from, the more benefit the hiring firm is likely to receive through the productivity gap. Equation 2.5 does not rule out the possibility of the hiring firm being exposed to beneficial knowledge from less productive firms. If hiring new workers is in general beneficial to a firm's exposure to knowledge, we will see this effect through the extensive margin, the hiring intensity terms.

All new hires are classified into one of the following sources ($\mathcal{S}$): (i) new workers for whom we have not observed any work history (e.g. new college graduates, new immigrants, etc); (ii) hires from firms outside of the scope of productivity analysis (i.e. hires from non-market or not private-for-profit firms); (iii) hires from very small firms for which the measure of productivity is likely to be particularly noisy (defined as less than five full-time equivalent workers); (iv) hires from private-for-profit firms within the scope of analysis but are missing some of the data required to compute productivity; and (v) hires from private-for-profit firms which are in scope and for which we have the data required to construct productivity gap measures. This latter group is the source of the productivity gap measures used to examine the intensive margin of knowledge spillover.

Disaggregated Productivity Gaps

Not all knowledge is likely to be equally useful to the hiring firm. For example, some knowledge carried by new employees may already be known by the firm, or the firm may have superior knowledge in that area already. In most of the analysis to follow, it will be appropriate to disaggregate the productivity gap into different productivity gaps for sub-groups of hires. This will allow us to estimate differences in the extent of knowledge spillover from each sub-group. For example, rather than use the aggregate productivity gap give in (2.5), the model for most of the analysis will use separate productivity gaps for hires from *more* and *less* productive firms.

Divide the set of new hires for which we can observe the productivity of the previous employer, $\mathcal{N}_{i,t-1}$, into two mutually exclusive sets $\mathcal{N}^M_{i,t-1}$ and $\mathcal{N}^L_{i,t-1}$ such that

$$\mathcal{N}^M_{i,t-1} \equiv \left\{ n \in \mathcal{N}_{i,t-1} : \ln(A_{n,\tau(n)}) - \ln(A_{i,t-1}) \geq 0 \right\}, \qquad \text{all } i, t, \tag{2.6}$$

represents the hires from more productive firms, and

$$\mathcal{N}^L_{i,t-1} \equiv \left\{ n \in \mathcal{N}_{i,t-1} : \ln(A_{n,\tau(n)}) - \ln(A_{i,t-1}) < 0 \right\}, \qquad \text{all } i, t, \tag{2.7}$$

represents the hires from less productive firms.

Using these two new subsets, the firm's exposure to knowledge from new hires can be written as

$$\begin{aligned} \text{Exposure}_{i,t} &= \beta_M \frac{\sum_{n \in \mathcal{N}^M_{i,t-1}} [\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\ &\quad + \beta_L \frac{\sum_{n \in \mathcal{N}^L_{i,t-1}} [\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\ &\quad + \sum_{s \in \mathcal{S}_{i,t-1}} \lambda_s \frac{H_{i,s,t-1}}{L_{i,t-1}}. \end{aligned} \tag{2.8}$$

It is important to note that the sum of the productivity differences in the first term of (2.8) is always positive and the second summation is always negative.

MFP and Measuring the Productivity Gap

As discussed by Fabling and Maré (2015b), it is not possible to directly compare the level of MFP between firms in different industries as MFP is measured relative to the average productivity in the industry. As a result, when hiring workers from another industry, the measures of the productivity gap are potentially biases because they excludes the difference between the average level of productivity between different industries.

Theoretically, no bias will exist when the knowledge that creates differences in average productivity between industries cannot be utilized by the hiring firm, but the part of the a firm's productivity that makes the worker's previous employer a highly productive firm in its own industry can. For example, if the hiring firm is unable to utilize the types of capital that makes the other industry more productive on average, but it is able to utilize the superior management practices that made the new worker's previous employer highly productive within its own industry.

Appendix B.2 looks at the potential for this bias to affect the results. The results there indicate that not accounting for the difference in average industry productivity does not appear to have a significant effect on the estimation results.

2.5.4 Theoretical Predictions

The aim of this paper is to investigate how firms benefit from the knowledge of new workers. The empirical analysis to follow will be compared to the predictions made by the hypothesis that labor acts as a channel for *knowledge spillover*. The second hypothesis considered is that a worker's own knowledge contributed to an *unmeasured worker quality* component. These predictions are discussed in more detail below. Table 2.1 provides a summary of the predictions from each hypothesis for the coefficients in the baseline model described by (2.8) and (2.4).

Table 2.1: Summary of theoretical predictions

	Parameter predictions for:		
	Productivity gap		Hiring intensity
Channel	β_M	β_L	λ_M & λ_L
Knowledge spillover	> 0	≈ 0	possibly $\lambda > 0$
Unmeasured worker quality	> 0	$\approx \beta_M$	

Notes: The parameter in this table relate to the model given by (2.8) and (2.4).

Knowledge Spillover

It has long been proposed in the literature that labor mobility may act as a channel for knowledge spillover between firms (see Glass and Saggi 2002 and Fosfuri et al. 1998 as examples). According to this effect, workers absorb some of the productive knowledge and ideas of their current employer while working on the job. Because knowledge is a non-rival good and not all knowledge can be protected, when a firm hires a worker with experience at another firm, it not only hires more labor input but also a stock of new ideas and knowledge. These ideas can be implemented by the hiring firm to augment their current production process.

Assuming adjustment costs are low enough and hiring firms have sufficient capacity to absorb the new knowledge, if labor mobility acts as a channel for productive knowledge spillover then hiring new workers from more productive firms should increase the hiring firm's stock of productive knowledge and hence productivity. Furthermore, the size of the productivity gain at the hiring firm should be positively correlated with the productivity of the new worker's previous employer. In the context of the model this effect would imply a positive coefficient for the productivity gap related to hires from more productive firms, i.e. $\beta_M > 0$.

Stoyanov and Zubanov (2012) argue that because firms are able to freely disregard any new knowledge that is less productive than the firm's current knowledge (e.g. a less efficient production technique), productive knowledge spillover from less productive firms should have very little impact on the hiring firm's performance. Therefore we should expect the productivity gap related to hires from less productive firms to have no effect, i.e. $\beta_L \approx 0$.

The amount of knowledge that workers are able to absorb and transmit is likely to be related to characteristics of the workers such as education, job type, or tenure. The ability of firms to absorb and implement new ideas may also be a factor is the spillover of knowledge. Therefore, we would expect to find larger spillover effects when the firm or worker have characteristics that could plausibility improve the ability of either to diffuse knowledge between firms.

Finally, if hiring firms are able to select workers for their knowledge when hiring, it is possible that we may not see any correlation between the productivity of the new worker's previous employer and the productivity gain at the hiring firm ($\beta_M \approx \beta_L \approx 0$). However, we would still expect that hiring intensity would still be correlated with the amount of knowledge spillover, and therefore the productivity gain at the hiring firm. Therefore, finding $\lambda_s > 0$ could also be considered consistent with the knowledge spillover channel.

Signal of Unmeasured Worker Quality

The regression model includes a control for the change in quality of the average worker ($\Delta Q_{j,t}$) which will be influenced by the quality of new workers arriving at the firm. As discussed in the next section, the measure of worker quality used in this analysis is derived from the worker's observed earnings across all jobs. However, there may be aspects of worker quality that are not capture by a workers wage. For example, labor market frictions or monopolistic power for the firm will mean that a worker's wage may not accurately reflect the worker's marginal product of labor. Therefore, there may be some unmeasured component of worker quality that the data does not account for.

If there is some unmeasured component of worker quality, there are multiple ways in which one might expect it to be correlation with firm productivity. One possibility is that high productivity firms are able to better screen and hire new candidates that are of high quality. Such a process would produce positive assortative matching between skilled workers and productive firms in the style of the work by Becker (1973).[7] Alternatively, productive firms might have better quality workers because they provide better on-the-job training or facilitate within-firm learning spillovers generated through interacting with the higher quality workforce already employed by the firm (see Nix 2015).

7. Most of the empirical work exploring firm-worker assortative matching has used two-way fixed effects regressions on wage data (see Abowd et al. 2004 as an example). However, this work does not measure firm productivity data directly.

When the previous employer's productivity provides a signal of a new hire's unmeasured quality, then hiring from more productive firms should raise the unmeasured quality of employed workers (and the hiring firms productivity), and hiring from less productive firms should lower the unmeasured quality of employed workers. I.e. $\beta_M > 0$ and $\beta_L > 0$. Furthermore, the magnitudes of the coefficients β_M and β_L should be approximately equal as hiring from symmetrically more or less productive firms should have a similar sized effect on the firm (but with opposite signs).[8]

2.6 Data

Information regarding firms comes from the Longitudinal Business Database (LBD) which combines a range of survey and administrative data sources for all economically significant businesses in New Zealand.[9] Information on employees comes from the Integrated Data Infrastructure (IDI) which links employers to employees via Pay-As-You-Earn (PAYE) tax records for each job, and also contains a wide range of other survey and administrative data sources on individuals linked by anonymized individual identification numbers.

The main analysis is conducted at the firm-year level. Firm-level financial year data is mapped to the nearest tax year ending March. The sample period for the analysis is from 2001 to 2013. Below is a summary of the data and key variables used in the paper.

8. The unmeasured worker quality channel may be sensitive to selection bias if the type of workers who select to leave more and less productive firms is not the same. Table 2.4 indicate that the average worker leaving more productive firms and the average worker leaving less productive firms both tend to be drawn from the lower part of the firm's earnings distribution, and move to similar rankings within the hiring firm (on average). This suggest that selection bias may not be a significant concern.

9. The term 'economically significant' encompasses firms that meet *at least one* of the following criteria: (i) More than $30,000 annual GST expenses or sales; (ii) more than three paid employees; (iii) in a GST exempt industry; (iv) part of a Business Register group of firms with ownership links; (v) a new GST registered firm. For more information on firm data in the LBD see Fabling and Sanderson (2016).

2.6.1 Firm Data

The unit of measurement for a firm is a Permanent Enterprise (PENT), as defined and developed by Fabling (2011). The PENT identifier is based on the firm identifier in the LBD, and corrects for certain events such as the change in the legal status of a firm. The scope of this analysis is restricted to private-for-profit businesses within the measured sector identified by Statistics New Zealand.[10] Only for these types of businesses do we believe that revenue and cost data will provide a suitable indicator for productivity.

For the main line of analysis, the MFP of firms is measured using both a Cobb-Douglas and trans-log production function using the estimates derived by Fabling and Maré (2015b). The MFP measures are estimated on annual data reflecting the fact that the survey and tax information on revenue and expenditures is only available at this frequency. The parameters of each production function is also allowed to vary across industries. In total, 39 separate industry classifications are used to cover all firms in the measured sector, similar in detail to level 3 of the ANZSIC06 New Zealand Standard Industrial Output Categories (NZSIOC).

In addition to the MFP measures of productivity, we also consider a measure of labor productivity computed as the (real) value added per worker. The measures of real output, materials, and labor used here are taken from the same LBD sources as those used by Fabling and Maré (2015b) to compute the MFP measures of productivity.[11]

According to Fabling and Maré (2015b), there are an average of 353,766 PENTs per year in the LBD with positive employment. Of these around 83 percent (292,978) are in the measured sector. Of the PENTs in the measured sector, around 32 percent are excluded from

10. Private-for-profit businesses broadly covers private producer enterprises, central and local government enterprises (i.e. trading departments of the government and State-Owned Enterprises), and private financial institutions. Notable exclusions include private households (including private production), government administration and defense, and private financial businesses. See Fabling and Sanderson (2016) for more details.
The measured sector is defined by Statistics New Zealand as "industries that mainly contain enterprises that are market producers. This means they sell their products for economically significant prices that affect the quantity that consumers are willing to purchase".

11. Appendix A provides further summary statistic information on the firm-level data.

our sample because they lack the necessary production information to estimate productivity. Finally, the productivity of very small firms is likely to be imprecisely measured, while measures of worker turnover in small firms can be both lumpy and extreme. Therefore, the scope of analysis is further restricted to only consider productivity growth for firms that employ an average of at least ten full time employees over the year. For the construction of the productivity gap, we allow the firm size of the worker's previous employer to be as low as an average of five full time equivalent workers.[12]

2.6.2 Worker Data

The worker data is used for two main purposes. First, to construct a measure of the average worker quality for each firm, and second, to map the transitions of workers between firms in order to construct the productivity gap measures.

Worker quality

The measure of average observed worker quality for each firm ($Q_{i,t}$) is computed weighting each worker by their contribution of total full-time equivalent (FTE) labor for the firm. The measure of individual worker quality/human capital is constructed following the approach of Hyslop and Maré (2009) who utilize the two-way fixed effects regressions on wage data developed by Abowd et al. (1999).

A worker's observed quality is given by the contribution of the worker fixed effect and the vector of worker-level observable characteristics to the worker's log wage, effectively

12. As shown in Table A.2 of Appendix A, the distribution of firm size within New Zealand is heavily dominated by very small firms, matching the predictions from Zipf's law. Raising the minimum firm size from an average of one FTE worker to ten FTE workers results in dropping around 90 percent of the PENT-years in the sample. The results of the regressions do not appear to be overly sensitive to the choice of minimum firm size.
The choice of different minimum firm sizes for the hiring firm and the previous employer is motivated by the fact that the analysis is not concerned with lumpy changes to firm size at the previous employer, only at the hiring firm. Therefore, by lowering the minimum firm size for the worker's previous firm when constructing the productivity gap, we can capture more of the labor flows in the economy in the measure of the productivity gap. The results do not differ much if the minimum firm size for the worker's previous employer is raised to ten.

stripping out the firm fixed effect and idiosyncratic error term.[13] This captures observable demographic characteristics alongside time-invariant characteristics including occupation, education and skill, and relies on an assumption that workers are fairly compensated for the value they bring to their employers.

The IDI only has information on hours worked for a minority of employees. Therefore we measure the labour supplied by each worker using the full-time equivalent (FTE) estimates developed by Fabling and Maré (2015a). This approach uses information on the worker's monthly income to estimate their labor supply, taking into account information like the statutory minimum wage, the number of jobs worked by the worker in a month, and the worker's income in adjacent months. One limitation of this method is that it is likely to over-estimate the labor input for some workers such as part-time workers who are highly paid.

Worker Transitions Between Firms

There are two issues that need to be addressed in mapping the transition of workers between firms in the data. First, when a new worker previously worked at multiple jobs, which firm(s) does the worker bring knowledge from? Second, because firm productivity is observed at the annual frequency, and worker transitions at the monthly frequency, what level of productivity knowledge exists at the previous employer in the month the worker leaves, and what level of productivity knowledge exists at the hiring firm in the month the worker arrives?

When worker have multiple jobs, it is assumed that the productive knowledge a new worker brings to the hiring firm comes from a single source, referred to as their "main job". A worker's main job is the one that pays the worker the most. Workers pay will be correlated with the time at that particular job and the worker's position within the firm's hierarchy.

13. The results do not differ significantly if worker quality is instead measured by only the worker fixed effect or the workers wage less the firm fixed effect. This suggests that the productivity gap and hiring intensities in the regression are not proxying for any worker-firm match quality that can be measured through the worker's wage.

Both of these factors are expected to give them more opportunities to acquire new knowledge.

The main job is determined as follows. If the worker is employed at multiple firms in the three months prior to starting their new job, the previous main job is the one from which the worker received the highest real (CPI adjusted) monthly income, for a full month's work, during this three-month window.[14]

If a new worker did not previously work at any job in the quarter before starting at their new firm, the employment history of the worker is traced back in time to the last month in which they were employed for the full month and the main job is determined from the jobs worked in that month. The analysis of does not make any allowance for depreciation of the worker's stock of knowledge and skills during jobless spells.

Because firm productivity is only observed annually, we must also address how to determine the firm's productivity in the month the worker leaves or joins a firm. We assume that if a worker leaves their previous employer in the first six months of that employer's financial year, it is assumed that the worker takes with them the productivity knowledge of the employer in the *previous* financial year. They do not observe/learn the firm's productivity knowledge for the current year because either it takes time for the worker to learn the new knowledge implemented this year, or the firm doesn't implement new productivity changes until part way through the year, after the employee has left. If the worker leaves in the last six months of their employer's financial year, it is assumed that the worker's productivity knowledge is based on the firm's productivity level for the current year. In the same way, when the worker starts at their new firm, the hiring firm's productivity level is based on annual productivity for the year six months prior to the worker's start month.

14. The reason only months in which workers are employed for the full month are considered here is that the income for months in which the worker begins/ends a job are imprecisely measured. For example, the paying out of any outstanding annual leave in the final month will bias upwards the workers income and not accurately reflect the work done that month.

2.6.3 Summary Statistics

Table 2.2 describes the firm-year characteristics of the private-for-profit firms in the sample. Across all the firms in the sample, the average size of the productivity gaps associated with hiring from more and less productive firms are similar (0.064 vs 0.060), leading to an aggregate productivity gap close to zero (0.005). This suggests that as a result of new hires, the average worker at the average firm has a value add last year that is 0.5 percent higher than the firm's value add last period. However, there is significant variation in the knowledge exposure measures for different firms as represented by the large standard deviation of the productivity gaps. Primarily this is due to the lumpy nature of the number of new hires each year, especially for smaller firms.

The distribution of labor across firms is highly skewed. In the sample of firms with an average of more than 10 FTE workers, the average firm uses the equivalent of around 56 FTE employees, while the median firm employees the equivalent of around 18 FTE employees on average across the year. The average firm also features a large amount of labor churn. On average, new workers supply just under 20 percent of the FTE labor units used by a firm each year. Workers who will leave the firm sometime during the current year supply on average around 17 percent of the firm's labor. This contributions to an excess turnover rate of around 50 percent.[15]

The average firm hires around 22 new employees each year, and the overwhelming majority of firm-years feature the firm employing at least one new worker. Around 49 percent of new workers come from other PFP firms that we observe productivity data from in the data set. In addition, 13 percent of all new hires by the average firm are from PFP firms within the same industry, and around 20 percent of new hires are from more productive PFP firms.

15. Excess turnover is computed as

$$\text{Excess turnover} = \frac{\text{starts} + \text{exits} - |\text{net change}|}{(FTE_t + FTE_{t-1})/2}$$

where FTE is the number of full time equivalent units of labor in the final month of the firm's financial year.

Table 2.2: Summary statistics at the firm-year level (Value-added per worker)

	Firms in sample ($FTE \geq 10$)			**Firms that hire new workers**			**Firms that hire from more productive firms**			**Firms that do not hire**		
Variable	**Mean**	**Median**	**S.D.**	**Mean**	**Median**	**S.D.**	**Mean**	**Median**	**S.D.**	**Mean**	**Median**	**S.D.**
Labor productivity												
log V.A. per worker	11.102	11.094	N.A.	11.101	11.093	N.A.	10.962	10.983	N.A.	11.176	11.165	N.A.
Growth rate V.A. per worker (%)	-0.004	0.000	0.432	-0.003	0.001	0.432	-0.003	0.001	0.457	-0.040	-0.019	0.388
Productivity gap												
Aggregate gap	0.005	0	0.208	0.005	0	0.209	0.041	0.015	0.228	0	0	0
More prod. firms gap	0.064	0.015	0.164	0.065	0.016	0.165	0.100	0.046	0.196	0	0	0
Less prod. firms gap	-0.060	-0.022	0.123	-0.061	-0.023	0.124	-0.059	-0.027	0.108	0	0	0
Labor force												
Total FTE units of labor	56.230	17.961	255.994	56.953	18.166	258.128	75.169	21.743	317.416	14.248	12.181	8.657
Share of FTE from new hires	0.194	0.155	0.169	0.198	0.157	0.169	0.218	0.180	0.162	0	0	0
Share of FTE from exiting workers	0.172	0.136	0.150	0.174	0.138	0.150	0.192	0.157	0.148	0.086	0.042	0.165
Excess (annual) turnover	0.514	0.457	0.329	0.522	0.462	0.325	0.594	0.538	0.330	0.019	0	0.054
New Hires												
No. of new employees	22.070	7	101.667	22.448	7	102.498	31.686	11	125.734	0	0	0
Share of hires from brand new workers	0.001	0	0.018	0.001	0	0.018	0.001	0	0.010	0	0	0
Share of hires from non-market	0.116	0.062	0.166	0.116	0.062	0.165	0.105	0.079	0.120	0	0	0
Share of hires from small firms (L<5)	0.288	0.250	0.232	0.288	0.250	0.231	0.260	0.250	0.171	0	0	0
Share of hires from missing prod. data	0.102	0.051	0.154	0.102	0.053	0.154	0.091	0.069	0.107	0	0	0
Share of hires from PFP	0.489	0.500	0.257	0.489	0.500	0.257	0.540	0.519	0.198	0	0	0
within same industry	0.131	0.061	0.180	0.131	0.062	0.180	0.148	0.105	0.170	0	0	0
More productive sources	0.205	0.167	0.219	0.205	0.167	0.219	0.305	0.250	0.202	0	0	0
Obs.	126048			124146			80700			1902		

Notes: Summary statistics based on the sample of firm-year observations in the data set. FTE refers to Full Time Equivalent units of labor (1 FTE = 1 worker per year). Shares of hires are computed as the number of hires from the subgroup relative to the total number of new hires for that firm-year. N.A. denotes values that have been censored in accordance with Statistics New Zealand's confidentiality guidelines. PFP denotes Private For Profit firms (those for which we have productivity data). 'Firms that hire from more productive firms' denotes any firm that hires at least one worker from a more productive firm during that year.

Table 2.2 also describes the characteristics of the subsets of firms that hire new workers, firms that hire at least one new worker from a more productive firm, and firms that do not hire new workers. Firms that hire at least one worker tend to have slightly lower productivity than firms that do not, but are also significantly larger in terms of labor force size, and have higher rates of labor market churn.

Firms do not hire new workers randomly and there is often a lot of selection (on both sides) when forming a new employment match. Table 2.3 shows where firms in each productivity decile source their new workers from. Remarkably, the share of new hires from each source are very similar for firms in all of the productivity deciles. The largest single source of new employees for firms in each productivity decile are from other firms with less than five employees.

Table 2.3: Worker transitions — Value-added per worker

	Source of new employee hires													
Hiring firm's	**PFP productivity decile**										**New**	**Non**	**Firms with**	**PFP miss.**
prod. decile	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Arrivals**	**Market**	**L<5**	**data**
1	0.05	0.08	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.00	0.16	0.30	0.08
2	0.05	0.08	0.05	0.05	0.04	0.04	0.04	0.04	0.03	0.03	0.00	0.16	0.31	0.08
3	0.04	0.07	0.06	0.06	0.04	0.04	0.03	0.03	0.03	0.03	0.00	0.16	0.32	0.08
4	0.04	0.06	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.00	0.14	0.33	0.08
5	0.04	0.06	0.05	0.05	0.05	0.04	0.04	0.05	0.04	0.04	0.00	0.14	0.32	0.08
6	0.04	0.06	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.04	0.00	0.13	0.33	0.08
7	0.04	0.06	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.00	0.13	0.32	0.08
8	0.04	0.06	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.06	0.00	0.13	0.32	0.08
9	0.04	0.05	0.03	0.03	0.04	0.04	0.04	0.06	0.06	0.08	0.00	0.13	0.31	0.08
10	0.04	0.05	0.03	0.03	0.03	0.03	0.04	0.05	0.06	0.13	0.00	0.14	0.28	0.09

Notes: Each cell shows the fraction of total hires made by all firms in each productivity decile (row) from each source (column). Sources are denoted by either their productivity decile (if the data is available), or are classified as out of scope due to the worker never being observed before (new arrivals), the sending firm being either non-market, the sending firm being a private for profit firm but too small for the sample, or there is missing productivity data for the PFP firm.
For example cell (1,1) states that firms in the lowest productivity decile hire 5 percent of their new hires from other firms in the lowest decile. Each row sums to one. Cells are shaded based upon the fraction of hires, with darker shades corresponding to a higher fraction of total hires. Deciles correspond to the firm's productivity ranking within each year, with decile 10 referring to the most productive firms.

In terms of hiring from other PFP firms for which we are able to estimate the productivity level, firms do marginally favor sourcing new workers from similar productivity deciles. However, the size of this bias is very small, and even firms in the lowest productivity decile still obtain about 4 percent of their new workers from firms in the top productivity decile. This suggests that the labor market is not segmented by firm productivity, and even the least productive firms still have fairly equal access to workers from the most productive firms.

Table 2.4 summarizes the key worker-level characteristics of new hires relative to different groups of workers. Panel A of the table shows the characteristics of new workers relative to the average incumbent worker in the hiring firm, one month after hiring. The average new worker earns an FTE income that is roughly 85 percent of the average incumbent worker at the hiring firm. Workers sourced from more productive firms tend to earn marginally more than workers from less productive firms (86.2 vs 84.5 percent of the average incumbent's earnings). New workers tend to be younger (about 88 percent of the average age) than the average incumbent, and less skilled (around 87 percent of the worker quality of the average incumbent worker). New workers are also more likely on average to be multiple job-holders, working an average of 10 percent more jobs in the same month.

Panel B of Table 2.4 shows the characteristics of new workers relative to the average worker at their previous main job (in the month prior to them leaving). On average, workers who change jobs tend to earn around 86 percent of the average FTE pay, supply 88 percent of the average FTE units, and also be younger and less skilled than the average worker. These results do not differ dramatically whether we consider workers coming from more or less productive firms. It is likely that the general negative selection in workers who leave firms may reflect the proportionally higher job mobility by younger/junior employees.

Panel C, shows the worker's characteristics at their new job, relative to their last main job. A workers FTE-adjusted monthly earnings are around 13 percent higher in their new job, and they work supply around twice the FTE units of labor. This large increase in labor supply is primarily driven by part-time workers and those with multiple jobs transitioning to full-time jobs. The median worker supplies the same number of hours at their new job as they did at their previous job. The average new employee has an average of just over five months break between jobs, with a median break of zero months. Given that we de not model any depreciation of skill or knowledge for long employment breaks, such a short duration between jobs is desirable.

Overall, workers who move between firms tend to come from the lower half of the sending

Table 2.4: Summary statistics for new workers

	All new hires			**New hires from more productive firms**			**New hires from less productive firms**		
Variable	**Mean**	**Median**	**S.D.**	**Mean**	**Median**	**S.D.**	**Mean**	**Median**	**S.D.**
A) New worker's characteristics (at the hiring firm) relative to incumbent workers									
Real earnings percentile	0.450	0.407	0.309	0.422	0.369	0.303	0.464	0.438	0.305
FTE supplied relative to avg. incument	0.903	1.003	0.690	0.904	1.002	0.616	0.918	1.008	0.715
Age relative to avg. incumbent	0.889	0.825	0.343	0.899	0.835	0.343	0.860	0.793	0.332
Worker quality percentile	0.490	0.478	0.299	0.460	0.429	0.293	0.511	0.500	0.295
Number of jobs relative to avg. incumbent	1.141	0.975	0.533	1.135	0.971	0.535	1.130	0.975	0.486
Obs.	4094400			1154500			1335200		
B) New worker's characteristics (at last main job) relative to the workers who stays									
Real earnings percentile	0.450	0.407	0.309	0.422	0.369	0.303	0.805	0.713	0.612
FTE supplied relative to avg. stayer	0.920	1.001	1.456	0.980	1.015	0.939	0.883	1	0.858
Age relative to avg. stayer	0.879	0.813	0.345	0.889	0.825	0.345	0.850	0.783	0.334
Worker quality percentile	0.490	0.478	0.299	0.460	0.429	0.293	0.511	0.500	0.295
Number of jobs relative to avg. stayer	1.145	0.972	0.546	1.131	0.965	0.543	1.139	0.974	0.504
Obs.	4005200			1131200			1314900		
C) New worker's characteristics at their new job relative to their own characteristics at the last main job									
Real earning per FTE	1.119	1.025	0.494	1.064	1.002	0.459	0.464	0.438	0.305
FTE supplied: new job relative to old job	2.346	1	228.217	2.178	1	205.699	2.532	1	330.921
No. of months between jobs	5.484	0	13.167	4.823	0	11.572	4.630	0	11.539
Prob. working in same industry	0.226	0	0.418	0.284	0	0.451	0.275	0	0.447
Obs.	4202000			1180200			1367800		

Notes: Summary statistics are computed at the worker-month level. Percentiles refer to the percentile within the firm (e.g. 0.45 implies the new worker is above 45 percent of workers in the firm). Statistics that are reported as relative to the average are computed as a fraction relative to the average member of the control group (e.g. 0.5 implies that the new worker's characteristic is half that of the average control group member). Worker quality is defined in section 2.6.2. FTE denoted Full Time Equivalent measure of labor. Real earnings are computed controlling for FTEs supplied.

firm's pool of labor (in terms of earnings, age, and labor supplied), and they also tend to have a similar ranking in the firms that they join. If knowledge spillover or unmeasured worker quality is related to observable characteristics of the workers, this relationship is likely to bias our baseline results downwards.

2.7 Analysis

The regression analysis is carried out in three stages. Section 2.7.1 analyses the baseline model. Section 2.7.2 extends the baseline model by disagregating the productivity gaps further. Section 2.7.3 considers issues to do with robustness.[16]

2.7.1 Baseline Model

Table 2.5 presents the initial regression results starting with a model where changes in firm productivity are driven only by changes in the quality of the labor force (column 1) and building up to the baseline model with separate productivity gaps for hires from more and less productive firms, defined by (2.4) and (2.8), in the final column. All results in the table are computed using value-added per worker as the productivity measure.

The first specification in Table 2.5 shows that the change in the firm's productivity is significantly correlated with the change in the (measured) quality of the average worker within the firm (ΔQ_i). According to the estimation, improving the quality of the average worker by 1 percent would be associated with average increase in labor productivity of around 0.5 percentage points.

Adjusting the average quality of a firm's labour through hiring is likely to incur different costs to the firm than adjusting through firing or changing the hours of incumbent workers. For example, newly hired workers may take time to adjust and fit to the culture of the firm. This can mean the effect of adjusting average worker quality through new hirings on

16. Other regression specifications are presented in Appendix B.

Table 2.5: Initial regression results

	Aggregate $\Delta Q_{i,t}$	$\Delta Q_{i,t}$ decomp.	Add share of new hires: All new hires	Add share of new hires: more/less decomp.	Add productivity gap: Aggregate prod gap	Add productivity gap: more/less decomp.	Add prod lags: Aggregate prod gap	Add prod lags: more/less decomp.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta Q_{i,t}$	0.507***							
	(0.051)							
$\Delta Q_{i,t}$ due to (γ):								
New hires		0.416***	0.432***	0.380***	0.359***	0.385***	0.467***	0.479***
		(0.051)	(0.050)	(0.050)	(0.050)	(0.049)	(0.076)	(0.075)
Exiters		0.407***	0.414***	0.366***	0.348***	0.373***	0.456***	0.468***
		(0.051)	(0.050)	(0.050)	(0.050)	(0.049)	(0.076)	(0.075)
Incumbents		0.431***	0.449***	0.396***	0.376***	0.402***	0.481***	0.494***
		(0.051)	(0.050)	(0.050)	(0.050)	(0.049)	(0.076)	(0.075)
Hire intensity (λ):								
New entrants			0.101	0.134	0.129	0.103	-1.466	-1.397
			(0.344)	(0.337)	(0.329)	(0.325)	(1.391)	(1.391)
Out of scope firms			-0.015	-0.052***	-0.071***	-0.071***	-0.085***	-0.091***
			(0.018)	(0.018)	(0.019)	(0.019)	(0.028)	(0.029)
Small PFP firms			-0.057***	-0.072***	-0.081***	-0.076***	-0.025	-0.024
			(0.015)	(0.014)	(0.014)	(0.014)	(0.022)	(0.022)
PFP firms missing data			-0.107***	-0.112***	-0.093**	-0.108***	-0.116**	-0.127**
			(0.036)	(0.034)	(0.029)	(0.032)	(0.047)	(0.050)
Observed PFP firms			-0.034***		-0.048***		-0.060***	
			(0.009)		(0.011)		(0.015)	
More prod. PFP firms				0.205***		-0.241***		-0.200***
				(0.014)		(0.043)		(0.057)
Less prod. PFP firms				-0.294***		-0.118***		-0.117***
				(0.018)		(0.020)		(0.027)
Excess turnover:			0.051***	0.046***	0.042***	0.044***	0.042***	0.044***
			(0.008)	(0.008)	(0.008)	(0.008)	(0.012)	(0.012)
Productivity gap (β):								
Aggregate gap					0.354***		0.281***	
					(0.022)		(0.027)	
More prod. firms						0.585***		0.480***
						(0.069)		(0.098)
Less prod. firms						0.165***		0.153***
						(0.020)		(0.030)
$\Delta \ln A_{i,t-1}$							-0.073**	-0.038
							(0.029)	(0.026)
Includes:								
Industry-year F.E.	yes	yes	yes	yes	yes	yes	yes	yes
Lagged productivity	no	no	no	no	no	no	yes	yes
Parameter tests:								
$\Pr(\beta_M = \beta_L)$						0.000		0.001
$\Pr(\lambda_M = \lambda_L)$				0.000		0.020		0.237
$\Pr(\gamma_{new} = \gamma_{incmb})$		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Obs.	89592	89592	88062	88062	84885	88062	36291	37269

Notes: The dependent variable $\Delta \ln A_{i,j,t}$. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

productivity growth is different to other adjustment margins. In the remaining specifications in the table, the change in average worker quality is decomposed into the contributions from new hires, workers who leave the firm, and incumbent workers.[17]

The specification in the second column of Table 2.5 shows the effects of this decomposition. While the effect of new workers is statistically different from the effect of the quality of labor through exiters and incumbents (p-value=0.000), the practical difference is small. A one percent increase in the quality of workers due to new hires raises the firm's productivity growth by 0.416 percentage points on average, while a one percent increase due to incumbent workers would raise the firm's productivity growth by 0.431 percentage points on average. We also see similarities in the magnitudes between the effect of the quality of new workers relative to incumbent workers for the other specifications shown in Table 2.5. This suggests that costs (in terms of the firm's productivity) associated with on-boarding new workers are not significant relative to the costs associated with improving the firm's average quality of labor via changes to the mixture of hours worked by incumbent employees.

The specifications in the third and fourth columns of Table 2.5 add to the model the hiring intensities from various sources, as well as the control for excess turnover. The coefficients on these hiring intensity variables can be interpreted as the average percentage point change in next period's productivity from hiring one percent of the labor force from that particular source, holding all else equal. In general, the coefficients are negative and suggest that hiring one percent of the firm's labor force from any of the sources usually lowers firm productivity in the region of 0.1 percentage points. The coefficients on excess turnover are positive, but small. The coefficient estimates suggest that excess hires on the order of 10 percent of the

17. More formally, let N denote new workers that join the firm in year t, I denote incumbent workers who work for the firm in both years t and $t-1$, and X denote workers who exit the firm between years $t-1$ and t. The change in worker quality can be equivalently written as:

$$\Delta Q_{i,t} = s_{N,t}Q_{N,t} - s_{X,t-1}Q_{X,t-1} + s_{I,t}Q_{I,t} - s_{I,t-1}Q_{I,t-1}$$

where $s_{A,\tau}$ denotes the share of labor for workers of type A at time τ, and $Q_{A,\tau}$ denotes the average quality of workers of type A at time τ within the firm. In the context of Table 2.5, the contribution from new hires is given by $s_{N,t}Q_{N,t}$, the contribution from those who exit is given by $-s_{X,t-1}Q_{X,t-1}$, and the contribution from incumbents is given by $s_{I,t}Q_{I,t} - s_{I,t-1}Q_{I,t-1}$.

firm's FTE labor supply is associated with a 0.4 percentage point increase in the growth rate of productivity.

The decomposition of the hiring intensities from PFP firms into hires from more productive PFP firms and hires from less productive PFP firms in column four yields an interesting insight into the small coefficient for the aggregate hire intensity from PFP firms (-0.034) seen in column three. Hiring new workers from more productive PFP firms leads to an increase, on average, of the hiring firm's productivity growth. This increase is in the order of 0.2 percentage points when hiring one percent of the work force comes from more productive PFP firms. However, hiring one percent of all workers from less productive PFP firms is associated with a decrease, on average, in the order of around 0.3 percentage points. By hiring new workers from a mixture of more and less productive PFP firms, the productivity gains from hiring from more productive sources are offset by the productivity losses from hiring from less productive sources. As a result, the average effect from hiring from PFP firms is relatively small. It is very likely that a similar offsetting is occurring in hiring from other sources. However, because we cannot observe the productivity of firms in these other sources, it is not possible to say with certainty.

The regressions in columns five and six of Table 2.5 add to the model the productivity gap variables, completing the inclusion of our proxy measure for the change in the firm's stock of productive knowledge. When considering hires from PFP sources in aggregate (column five), the coefficient on the aggregate productivity gap (β_2) suggests that for a firm that has a hiring intensity (H/L) from PFP sources of 10 percent, raising the average productivity of the PFP firms that workers are sourced from by one percent would be associated with an average 0.35 percentage point increase in productivity growth for the hiring firm.[18] Column six shows that if we disaggregate the productivity gap into separate productivity gaps for

18. As an alternative interpretation of the coefficient, one could view the coefficient through the lens of the average worker's exposure to better productivity. Hiring from other PFP sources such that the average worker within the firm has previous productivity knowledge one percent greater than the hiring firms productivity will raise the productivity growth in the hiring firm by 0.35 percentage points on average.

hires more and less productive firms, the productivity gain (intensive margin) associated with the productivity gap for hiring from more productive firms is twice as large as the productivity loss associated with hiring from symmetrically less productive firms.

The final two specifications in Table 2.5 include the lagged productivity dynamics of the hiring firm and represent the baseline model. In practice, most of the coefficients are not significantly affected by the inclusion of productivity lags. The coefficients related to the change in worker quality are slightly larger, and the coefficients related to the productivity gaps are slightly smaller, but the differences are relatively small.

The fact that the coefficients on the productivity gaps for hires from both more and less productive firms are both positive and significant is consistent with an unmeasured worker quality channel (which predicts that hiring from even more productive firms should raise the unmeasured worker quality within the firm, and hiring from even less productivity firms should lower it). In addition, the fact that the coefficient on the productivity gap associated with hires from more productive firms is significantly larger than that on the productivity gap associated with hires from less productive firms is consistent with the productive knowledge spillover story.

If we were to assume that both the signal of unmeasured worker quality channel and the productive knowledge spillover channel were occurring simultaneously, this assumption would suggest that the size of the knowledge spillover premium for hiring from more productive firms would be equal to 0.33, the difference between the two coefficients (0.48-0.15). This implies that a little over two thirds of the improvement associated with the increase in the average source's productivity would be due to the knowledge spillover, and around one third would be due to improvements in the unmeasured worker quality.

Labor productivity, or value-added per worker, is only one possible measure of firm productivity. Table 2.6 compares the estimated key parameters from the baseline model using value-added per worker to the estimated values found using various MFP measures of firm productivity. The coefficients related to the productivity gaps for all productivity measures are positive and significant in magnitude. However, in the case of the Cobb-Douglas measure of MFP, the coefficient related to the productivity gap from less productive firms is around 40 percent larger than the coefficient related to the productivity gap from more productive firms (0.374 compared to 0.271). However this difference is not significantly different (p-value = 0.22). In the case of the trans-log based measure of productivity (whose specification nests the Cobb-Douglas), the coefficients on the two productivity gaps that are very similar, both economically, as well as statistically (p-value = 0.8).

Additionally, the coefficients related to the hire intensity (λ_s) from more and less productive firms are no longer significantly negative when using the MFP measures of productivity, and are very close to zero. In fact for the Cobb-Douglas measure of productivity, the coefficient on the hiring intensity from less productive firms is slightly positive. However, in general the hiring intensities do not seem to have a significant influence on productivity growth when using MFP to measure firm productivity.

For all productivity measures, even after controlling for observed worker quality, the coefficients on the productivity gaps are positive, implying that raising the productivity of the average PFP firms workers are sourced from leads to higher productivity growth on average. This finding supports the idea of an unmeasured worker quality channel (which predicts both coefficients should be positive and equal). This is especially true for the MFP measures of firm productivity where the coefficients on the productivity gaps relate to hires from more and less productive firms are not statistically different.[19]

Only the baseline model estimated using value-added per worker as the measure of firm productivity provides support the predictions of the knowledge spillover channel (which predicts a larger coefficient on the productivity gap associated with hires from more productive firms). One of the potential reasons why we see this support in the value-added productivity measure and not the MFP based measures is that the MFP measures of productivity control

19. Estimating the model seperately for each industry reveals that the baseline results are fairly consistent across the largest industries in the data set. Therefore, the results are not being biased by one particular industry.

Table 2.6: Baseline regression results for various productivity measures

	Value-added	**Cobb-Douglas**	**Trans-log**
Productivity gap, hires from (β):			
More prod. Firms	0.480***	0.271***	0.354***
	(0.098)	(0.065)	(0.068)
Less prod. Firms	0.153***	0.374***	0.374***
	(0.030)	(0.054)	(0.056)
Hire intensity (λ):			
More prod. firms	-0.200***	-0.012	-0.037*
	(0.057)	(0.028)	(0.021)
Less prod. Firms	-0.117***	0.047*	0.004
	(0.027)	(0.026)	(0.019)
$\Delta Q_{i,t}$ due to (γ):			
New hires	0.479***	0.105*	0.162***
	(0.075)	(0.062)	(0.049)
Exiters	0.468***	0.103*	0.159***
	(0.075)	(0.062)	(0.048)
Incumbents	0.494***	0.110*	0.166***
	(0.075)	(0.062)	(0.048)
Parameter tests:			
$\Pr(\beta_M = \beta_L)$	0.001	0.217	0.808
$\Pr(\lambda_M = \lambda_L)$	0.237	0.145	0.174
$\Pr(\gamma_{new} = \gamma_{incmb})$	0.000	0.037	0.026
Obs.	37269	28260	38037

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. When included in the regression, $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

for the use of capital and materials. If firms that hire workers from other firms with higher labor productivity end up increasing their own capital utilization, this hiring would appear as an increase in labor productivity, but not necessarily as an increase in MFP.[20] Therefore, if the larger coefficient on the productivity gap for hires from more productive firms seen in the value-added results does relate to a knowledge spillover channel, it is likely that the knowledge relates to production technology (the functional form of the production function, or how capital intensive the production process is), rather than the strict multi-factor/total-factor productivity of the firm.

To investigate this hypothesis further, the baseline model is re-estimated using the firm's ratio of capital to labor as dependent variables instead of firm productivity. Table 2.7 summarizes the key coefficients from this regressions. The coefficients related to the input intensity gap (the replacement for the productivity gap) for hires from more capital-intensive firms is around twice as large as the coefficient related to the input intensity gap for hires from less input-intensive firms. This suggests that there is an increase in input-intensity associated with hiring from firms that are more input-intensive, matching the pattern seen in the productivity gap coefficients for the value-added measure of firm productivity.[21]

To summarize the findings of the baseline regressions: while the baseline results do not imply causality or provide definitive conclusions regarding the channels through which new workers benefit hiring firms, they do point us towards the likely channels which would be consistent with the findings. For all of the firm productivity measures considered, both labor productivity and MFP measures, the coefficients on the productivity gaps for hires from more and less productive firms are both positive and significant. Therefore, all else equal, when

20. Another possibility is the fact that the MFP measures fail to capture the productivity level differences between industries. This issue is explored further in the following subsection.

21. Another possible driver of the differences between the value-added and MFP results is that MFP measures of productivity are constructed relative to an industry-year average. Hence, when constructing the productivity gap using MFP measures, we fail to capture any between-industry productivity differences, which the value-added measure of labor productivity would capture. This issue is explored further in Appendix B.2 by re-estimating the model on demeaned value-added data. The results do not differ significantly from the regular value-added results, suggesting that the demeaned nature of MFP measures is not the main driver of the differences between the results using value-added and MFP measures.

Table 2.7: Baseline results for capital-labor ratio measure

	Capital-Labor
Input intensity gap, hires from (β):	
More capital-intensive firms	0.047***
	(0.017)
Less capital-intensive firms	0.021
	(0.024)
Hire intensity (λ):	
More capital-intensive firms	0.071**
	(0.031)
Less capital-intensive firms	-0.182***
	(0.035)
$\Delta Q_{i,t}$ due to (γ):	
New hires	0.568***
	(0.075)
Exiters	0.558***
	(0.075)
Incumbents	0.608***
	(0.075)
Parameter tests:	
$\Pr(\beta_M = \beta_L)$	0.369
$\Pr(\lambda_M = \lambda_L)$	0.000
$\Pr(\gamma_{new} = \gamma_{incmb})$	0
Obs.	28260

Notes: The dependent variable in the regressions is the change in log capital-labor ratio ($\Delta \ln(K_{i,j,t}/L_{i,j,t})$). Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln \ln(K_{i,j,t-1}/L_{i,j,t-1})$ is instrumented for using $\ln(K_{i,j,t-2}/L_{i,j,t-2})$ in response to the presence of Nickell bias. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

hiring from other private for profit (PFP) firms, raising the average productivity of the firms workers are sourced from improves productivity growth next period. This pattern is consistent with the idea of an unmeasured worker quality component in which worker quality is unobserved by the econometrician but is signalled to the hiring firm through the productivity of the worker's previous employer.

In the specific case of labor productivity, the coefficient on the productivity gap for hires from more productive firms is more than twice as large as the coefficient on the productivity gap from less productive firms. This would suggest that there is a premium (over and above the unmeasured worker quality channel) from hiring from more productive firms, consistent with the predictions of a knowledge spillover channel. However, once we move to looking at firm productivity through the lens of MFP measures, this premium disappears. By estimating the model on the capital-labor ratio, we see that there is also an input-intensity gap premium associated with hiring from more input-intensive sources. This, combined with the MFP findings, suggests that if the productivity gap premium is being driven by a knowledge spillover effect, the knowledge specifically refers to knowledge regarding production technology (the functional form of the production function) rather than multi-factor productivity knowledge. In other words, firms are able to adopt more input-intensive production techniques using the knowledge of workers with more experience in these approaches.

2.7.2 Extensions to the Model

In this section the baseline model will be extended to consider how various firm and worker characteristics influence the productivity gap and hiring intensity coefficient estimates. These extensions are based on the predictions made by the various channels of worker benefit considered by the analysis, and will provide further checks on the strength of support for the channels found so far.

Industry-Specific Knowledge

If workers facilitate the spillover of knowledge between firms, not all knowledge that workers bring into the firm will be of equal value. The structure of the baseline model already allows for different effects from knowledge coming from more and less productive firms through the disaggregated productivity gaps. However, this dimension is not the only dimension along which the value of knowledge will differ for the hiring firm. For example, workers with knowledge that relates to the market in which the hiring firm operates or knowledge that is able to complement the hiring firm's current stock of productive knowledge are likely to have a greater effect on firm productivity than workers with other types of knowledge. Therefore, the knowledge spillover channel predicts that workers hired away from other firms within the same industry (whose knowledge should be more valuable to the hiring firm) would have a larger benefit for the hiring firm's productivity than workers hired away from firms in other industries.

To examine if this is the case, the productivity gaps (and hire intensities) related to hires from more and less productive PFP firms are further subdivided into two groups: hires from within the same industry, and hires from different industries, i.e.:

$$\begin{aligned}
\text{Exoposure}_{i,t} \quad = \quad & \sum_{\text{ind}\in\{\text{same,diff}\}} \beta_{M,\text{ind}} \frac{\sum_{n\in\mathcal{N}^M_{j,t-1}} \mathbb{D}_{\text{ind}}(n)[\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\
& + \sum_{\text{ind}\in\{\text{same,diff}\}} \beta_{L,\text{ind}} \frac{\sum_{n\in\mathcal{N}^L_{j,t-1}} \mathbb{D}_{\text{ind}}(n)[\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\
& + \sum_{s\in\mathcal{S}_{i,t-1}} \lambda_s \frac{H_{i,s,t-1}}{L_{i,t-1}}, \qquad (2.9)
\end{aligned}$$

where 'ind = same' denotes the hire is from the same industry, 'ind = diff' denotes the hire is from a different industry, $\mathbb{D}_{\text{ind}}(n)$ is a dummy variable based on the 'ind' classification. So when $\mathbb{D}_{\text{ind}}(n) = \mathbb{D}_{\text{same}}(n)$, $\mathbb{D}_{\text{same}}(n)$ takes on the value of 1 if worker n's previous main job was in the same industry (and a similar definition for the case when 'ind = diff'). Therefore,

$\beta_{M,\text{same}}$ denotes the effect of the productivity gap for hires from more productive firms within the same industry. In addition, the set of sources, $\mathcal{S}_{i,t-1}$, is expanded to include hires from the same and different industries who worked at more or less productive firms.

Table 2.8 presents the key results for estimating this extended version of the model for the three main productivity measures. There are 39 different industries within the data set (at roughly a 3-digit level of classification). With such a narrow definition of same industry, there is the possibility that the knowledge from other closely related industries might also be highly applicable, and this effect will be lost when aggregating hires from other relevant industries with those from less relevant industries.[22] Therefore, Table 2.8 also provides results for when the definition of same industry is based on industry groups aggregated to the 1-digit level (e.g. all of manufacturing industries are grouped together).

Table 2.8 shows that when value-added per worker is used to measure firm productivity, the coefficient on the productivity gap from workers from more productive firms within the same industry is nearly three times as large as the coefficient on the productivity gap from workers from more productive firms in other industries (the p-value is 0.03). The coefficients on the productivity gap from less productive firms are (i) significantly lower then the coefficients on the productivity gap from more productive firms and (ii) not significantly different between hires from the same and hires from different industries. Relative to the baseline results, this pattern in productivity gap coefficients supports the predictions of a productive knowledge spillover channel that productive knowledge from within a firm's own industry is more applicable to the hiring firm and provides a larger boost to firm productivity than productive knowledge from outside the industry. Less productive knowledge, whether from inside or outside the firm's industry, is less useful to the hiring firm, and will likely be discarded.

For both of the MFP measures considered, the coefficients related to the productivity

22. For example, the 'Sheep, beef cattle, and grain farming' industry and the 'Dairy cattle farming' industry appear as separate industries in the data at the 3-digit level, but likely share some common knowledge base.

Table 2.8: Regression results featuring between and within industry productivity gaps

	Value-added			Cobb-Douglas			Trans-log		
	Baseline	Productivity gaps by ind.		Baseline	Productivity gaps by ind.		Baseline	Productivity gaps by ind.	
		3-digit	1-digit		3-digit	1-digit		3-digit	1-digit
Prod. gap, hires from (β):									
More prod. firms	0.480***			0.271***			0.354***		
	(0.098)			(0.065)			(0.068)		
Within same ind.		1.016***	0.926***		0.234*	0.156		0.311**	0.287***
		(0.230)	(0.164)		(0.137)	(0.099)		(0.123)	(0.105)
From diff. ind.		0.367***	0.326***		0.296***	0.350***		0.373***	0.390***
		(0.123)	(0.121)		(0.088)	(0.100)		(0.090)	(0.097)
Less prod. firms	0.153***			0.374***			0.374***		
	(0.030)			(0.054)			(0.056)		
Within same ind.		0.188***	0.156***		0.565***	0.572***		0.278***	0.314***
		(0.068)	(0.052)		(0.142)	(0.110)		(0.103)	(0.079)
From diff. ind.		0.139***	0.152***		0.319***	0.286***		0.415***	0.418***
		(0.035)	(0.039)		(0.065)	(0.074)		(0.071)	(0.086)
Hire intensity (λ):									
More prod. firms	-0.200***			-0.012			-0.037*		
	(0.057)			(0.028)			(0.021)		
Within same ind.		-0.310***	-0.312***		0.036	0.042		0.008	0.002
		(0.075)	(0.064)		(0.042)	(0.035)		(0.028)	(0.028)
From diff. ind.		-0.168*	-0.154*		-0.037	-0.055		-0.059**	-0.066**
		(0.078)	(0.081)		(0.038)	(0.044)		(0.028)	(0.031)
Less prod. firms	-0.117***			0.047*			0.004		
	(0.027)			(0.026)			(0.019)		
Within same ind.		-0.069	-0.099***		0.089*	0.088**		-0.001	-0.005
		(0.040)	(0.035)		(0.046)	(0.038)		(0.028)	(0.023)
From diff. ind.		-0.139***	-0.120***		0.031	0.028		0.004	0.010
		(0.033)	(0.037)		(0.032)	(0.037)		(0.025)	(0.029)
Parameter tests:									
$\Pr(\beta_M = \beta_L)$	0.001			0.217			0.808		
$\Pr(\beta_{M,\text{same}} = \beta_{L,\text{same}})$		0.001	0.000		0.080	0.004		0.825	0.832
$\Pr(\beta_{M,\text{diff}} = \beta_{L,\text{diff}})$		0.064	0.159		0.833	0.600		0.712	0.824
$\Pr(\beta_{M,\text{same}} = \beta_{M,\text{diff}})$		0.029	0.005		0.728	0.214		0.724	0.517
$\Pr(\beta_{L,\text{same}} = \beta_{L,\text{diff}})$		0.531	0.955		0.138	0.048		0.302	0.406
$\Pr(\lambda_M = \lambda_L)$	0.237			0.145			0.174		
$\Pr(\lambda_{M,\text{same}} = \lambda_{L,\text{same}})$		0.008	0.006		0.418	0.389		0.846	0.851
$\Pr(\lambda_{M,\text{diff}} = \lambda_{L,\text{diff}})$		0.748	0.733		0.202	0.177		0.111	0.095
$\Pr(\lambda_{M,\text{same}} = \lambda_{M,\text{diff}})$		0.228	0.117		0.228	0.104		0.124	0.117
$\Pr(\lambda_{L,\text{same}} = \lambda_{L,\text{diff}})$		0.169	0.659		0.305	0.270		0.895	0.694
Obs.	37269	37269	37269	28260	28260	28260	38037	38037	38037

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$, where the measure of productivity differs by column. The 3-digit classification refers to the level of industry classification used by Fabling and Maré (2015b) which is very similar to the level 3 ANZSIC06 categories. The 1-digit classification refers to the level 1 ANZSIC06 categories. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

gaps from hires in the same industry are not significantly different from those related to hires from different industries (or the baseline results that do not distinguish between industries). This broadly lines up with the prediction from the unmeasured worker quality channel that the benefit to the hiring firm is unlikely be affected by the industry the worker previously worked in, given that it is hard to motivate how there will be a systemic difference in the ability of firms to screen or train workers within and between industries.[23]

The remaining parameters in the model are generally no significantly affected by the distinction between hires from within or between industries.Most notably, the coefficients related to the hire intensities (shown in Table 2.8), which capture the effect of hiring intensity from the various sources, do not differ significantly with hiring from the same or different industries.[24]

Tenure

Workers who have been at their previous employer for a long period of time are more likely to have acquired knowledge about the productivity practices of that firm, more likely to have received on-the-job training, and more likely to have had a good fit with their employer.

According to the knowledge spillover channel, workers who have a longer tenure at their previous employer should have more opportunities to observe and learn what makes their employer productive. As a result, the amount of knowledge spillover from more productive firms should be positively correlated with the length of tenure at the previous firm. Workers from less productive firms are not likely to be transmitters of knowledge between firms (since their knowledge is likely to be inferior), and productivity in the hiring firm should hence not be affected by the tenure length of these workers.

23. An exception to this prediction would be if unmeasured worker quality was related to on-the-job training, and workers received training that was industry specific. In such a case, we would expect to see some differences in the coefficients related to the same and different industries.

24. Because the coefficient on the productivity gap of hires from different industries are similar to those of hires from the same industry (which does not suffer from bias from average MFP differences across industries, it is unlikely that using productivity differences calculated from MFP suffer from significant bias effects.

The predictions of the unmeasured worker quality channel in relation to worker tenure depend upon what mechanism underlies the positive assortative matching between firms and workers. If productive firms are providing better quality training to their workers (making them more productive at future employers), we would expect to see larger productivity gains for longer tenured workers. However, if productive firms are simply better at screening workers (making the previous employer a signal for the worker's innate productivity), then their tenure at the previous employer is less likely to have a dramatic effect on the productivity gains to the hiring firm.

To explore these ideas further, the productivity gaps and hiring intensities in the baseline model are sub-divided into workers with long tenure, and workers with short tenure. More formally, the change in the firm's knowledge in the baseline model now takes on the form

$$\begin{aligned}
\text{Exposure}_{i,t} \quad = \quad & \sum_{\text{tenure}\in\left\{\substack{\text{long,}\\ \text{short}}\right\}} \beta_{M,\text{tenure}} \frac{\sum\limits_{n\in\mathcal{N}_{j,t-1}^{M}} \mathbb{D}_{\text{tenure}}(n)[\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\
& + \sum_{\text{tenure}\in\left\{\substack{\text{long,}\\ \text{short}}\right\}} \beta_{L,\text{tenure}} \frac{\sum\limits_{n\in\mathcal{N}_{i,t-1}^{L}} \mathbb{D}_{\text{tenure}}(n)[\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\
& + \sum_{s\in\mathcal{S}_{i,t-1}} \lambda_s \frac{H_{i,s,t-1}}{L_{i,t-1}}, \qquad (2.10)
\end{aligned}$$

where 'tenure = long' denotes workers with long tenure, 'tenure = short' denotes workers with short tenure, $\mathbb{D}_{\text{tenure}}(n)$ is a dummy variable that takes on the value 1 if worker n has tenure length given by 'tenure' in their previous main job, and 0 otherwise. Therefore $\beta_{M,\text{tenure}}$ denotes the effect of the productivity gap associated with hires of workers from more productive firms who have long tenure at that firm.

One potential issue with the analysis of tenure described above is that it doesn't control for time spent at the hiring firm. If workers with long tenure at their previous main job are not spending enough time at the hiring firm, they may be unable to have much of an effect on the hiring firm's productivity. So as a further extension to the analysis, a second definition of 'long tenure workers' is also considered. In this alternative definition, a long tenured worker is one who has had a tenure of at least 12 months in their previous main job before being hired, and who also spend at least 12 months employed in the hiring firm (all new hires who fail to meet both of these conditions are grouped together into the short tenured group). Imposing this extra tenure requirement on the time spent at the hiring firm ensures that workers hired in the previous period remain employed at the firm long enough to influence the firm's production in the current period (since the productivity gap is based on hires in year $t-1$, and the dependent variable is productivity growth in year t).

For each productivity measure, Table 2.9 presents three columns of results. The first column is the baseline regression results seen previously. The second column decomposes the productivity gaps and hire intensities based on the worker's length of tenure at their previous firm. The third column uses the alternative definition of long tenured workers based on their time at both the previous and the hiring firms.

Turning first to the definition of long tenured workers based solely on tenure length at their previous employer (the sending firm), for all of the productivity measures in Table 2.9, the point estimates of the coefficients related to both the productivity gaps from more productive firms and also the coefficients on the hiring intensities are broadly similar across short and long term workers). However, the coefficients on the productivity gaps related to hires from less productive firms do show some systematic differences from the baseline across all productivity measures (although not significantly different). For all productivity measures, the coefficient on the productivity gaps related to hires from less productive firms is larger for longer tenured workers than short tenured workers.

When the 12-month tenure requirement for long tenured workers is imposed at both the sending and hiring firm, the coefficient on the productivity gap for long tenured hires from more productive firms rises from 0.375 to 0.754 when using value-added per worker

Table 2.9: Effects of considering worker tenure

	Value-added			Cobb-Douglas			Trans-log		
	Baseline	Tenure at:		Baseline	Tenure at:		Baseline	Tenure at:	
		Sending	Sending & hiring		Sending	Sending & hiring		Sending	Sending & hiring
Prod. gap, hires from (β):									
More prod. firms	0.480***			0.271***			0.354***		
	(0.098)			(0.065)			(0.068)		
With long tenure		0.375***	0.754***		0.331***	0.316*		0.325***	0.552**
		(0.138)	(0.280)		(0.109)	(0.190)		(0.116)	(0.229)
With short tenure		0.550***	0.419***		0.216***	0.254***		0.343***	0.292***
		(0.155)	(0.115)		(0.075)	(0.075)		(0.091)	(0.078)
Less prod. firms	0.153***			0.374***			0.374***		
	(0.030)			(0.054)			(0.056)		
With long tenure		0.187***	0.255***		0.599***	0.766***		0.569***	0.679***
		(0.053)	(0.066)		(0.114)	(0.161)		(0.087)	(0.130)
With short tenure		0.116**	0.099**		0.232***	0.288***		0.226***	0.281***
		(0.045)	(0.045)		(0.087)	(0.059)		(0.086)	(0.068)
Hire intensity (λ):									
More prod. firms	-0.200***			-0.012			-0.037*		
	(0.057)			(0.028)			(0.021)		
With long tenure		-0.142*	-0.247*		-0.022	0.010		-0.024	-0.053
		(0.073)	(0.131)		(0.041)	(0.060)		(0.032)	(0.058)
With short tenure		-0.242***	-0.190***		-0.005	-0.018		-0.041	-0.031
		(0.085)	(0.068)		(0.034)	(0.035)		(0.028)	(0.025)
Less prod. firms	-0.117***			0.047*			0.004		
	(0.027)			(0.026)			(0.019)		
With long tenure		-0.123***	-0.140**		0.067*	0.120**		0.041	0.072
		(0.043)	(0.056)		(0.041)	(0.057)		(0.030)	(0.048)
With short tenure		-0.111***	-0.125***		0.048	0.030		-0.023	-0.019
		(0.037)	(0.035)		(0.040)	(0.031)		(0.026)	(0.021)
Parameter tests:									
$\Pr(\beta_{M,\text{long}} = \beta_{L,\text{long}})$		0.200	0.085		0.094	0.074		0.097	0.630
$\Pr(\beta_{M,\text{short}} = \beta_{L,\text{short}})$		0.006	0.007		0.891	0.724		0.344	0.916
$\Pr(\beta_{M,\text{long}} = \beta_{M,\text{short}})$		0.445	0.297		0.421	0.773		0.915	0.316
$\Pr(\beta_{L,\text{long}} = \beta_{L,\text{short}})$		0.350	0.092		0.030	0.008		0.013	0.016
$\Pr(\lambda_{M,\text{long}} = \lambda_{L,\text{long}})$		0.835	0.449		0.133	0.185		0.149	0.094
$\Pr(\lambda_{M,\text{short}} = \lambda_{L,\text{short}})$		0.191	0.449		0.331	0.321		0.651	0.735
$\Pr(\lambda_{M,\text{long}} = \lambda_{M,\text{short}})$		0.383	0.713		0.753	0.702		0.705	0.750
$\Pr(\lambda_{L,\text{long}} = \lambda_{L,\text{short}})$		0.841	0.841		0.754	0.190		0.134	0.105
Obs.	37269	37269	37269	28260	28260	28260	38037	38037	38037

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$, where the measure of productivity differs by column. The cut off length for distinguishing between long and short tenure is equal to 12 months of previous employment at the respective firm. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

as the firm's productivity measure.[25] For hires from less productive firms, the increase in the productivity gap coefficient related to long tenured hires is relatively small (from 0.187 to 0.255). When MFP is used, there is a general increase in the coefficient on the productivity gap for long tenured workers hired from less productive firms (an increase of 0.167 for Cobb-Douglas and 0.11 for trans-log), and in the case of trans-log productivity, we also see an increase in the productivity gap coefficient related to long tenured workers from more productive firms (from 0.325 to 0.552). The remaining parameters in the model are generally not affected by the distinction between long and short tenured workers.

From the perspective of the knowledge spillover channel in the value-added data, the data suggests longer tenure at the sending firm is actually related to less, not more, gains in the productivity gap from more productivity firms. And in the case of hires from less productive firms, longer tenure at the sending firm is actually correlated with lower productivity growth at the hiring firm. Both of these contradict the predictions of the knowledge spillover channel (where longer tenure at the sending firm should be beneficial to the hiring firm). When tenure at both the sending and hiring firms are considered, we do see the expected larger coefficient on longer tenured hires from more productive firms relative to shorter tenured workers. However, the coefficient on the productivity gap of longer tenured workers from less productive firms still remains larger than for shorter tenured workers.

The unmeasured worker quality channel would suggest we might see longer tenured workers at the sending firm having more of a positive influence on the hiring firm's productivity when workers receive some form of on-the-job training. However, in the MFP we see little difference between the coefficients related to long and short tenured workers from more productive firms. In the case of workers hired from less productive firms, the productivity gap coefficient related to longer tenured workers suggests that for a given difference in the productivity between the sending and hiring firms, hiring workers with longer tenure at the

25. With the way long tenured workers are defined in this second set of results, the most relevant comparison of parameters to make is between long tenured workers under both definitions. This illustrates the effect of tenure at the hiring firm for these workers.

sending firm results in lower (not higher) productivity growth.

Table 2.10 repeats the above decomposition into long-tenured and short-tenured workers for the model estimated using the capital-labor ratio rather than productivity. The premium in the input intensity gap associated with hiring workers from more productive firms is driven predominantly by workers with more than 12 months tenure at their previous main employer.

The results for the capital-labor ratio are in line with what one would expect to see if there is a knowledge spillover from more input intensive firms. Having a longer tenure at more input-intensive firms would give the workers an opportunity to acquire more knowledge/experience with these more input-intensive production methods, and hence workers from more productive firms should be able to transmit more knowledge to the hiring firm relative to a worker with less experience. Tenure does not seem to affect the input-intensity gains much for hires from less input intensive firms, who are less likely to transmit new knowledge.

Overall, the relationship between worker tenure and the productivity gains from the productivity gaps do not support any of the channels considered. However, when looking at the capital-labor ratio, we do find support for tenure benefiting the input-intensity gains, and this result is consistent with the predictions of the knowledge spillover channel.

Worker's Skill Complementarity

The results of the baseline regressions using MFP measures are broadly in line with the predictions made by the unmeasured worker quality channel. Another prediction from this channel that can be tested is whether or not the coefficients on the productivity gaps vary with measured worker quality. Because the baseline model already controls for measured worker quality (through the regressor $\Delta Q_{i,t}$), the productivity gap relates to the potential productivity gains/losses to the hiring firm from the component of worker skill that is not already captured by the measure of worker quality derived from the worker's wage data. As a result, the unmeasured worker quality channel predicts that the effect of the productivity

Table 2.10: Effects of considering worker tenure — capital-labor ratio

	Baseline	**Tenure at:**	
		Sending	**Sending & hiring**
Capital intensity gap, hires from (β):			
More capital-intensive firms	0.047***		
	(0.017)		
With long tenure		0.138***	0.376***
		(0.038)	(0.095)
With short tenure		-0.011	0.004
		(0.033)	(0.019)
Less capital-intensive firms	0.021		
	(0.024)		
With long tenure		0.077*	0.110*
		(0.042)	(0.059)
With short tenure		0.071	0.004
		(0.050)	(0.028)
Hire intensity (λ):			
More capital-intensive firms	0.071**		
	(0.031)		
With long tenure		0.031	-0.187**
		(0.054)	(0.090)
With short tenure		0.071	0.101***
		(0.050)	(0.038)
Less capital-intensive firms	-0.182***		
	(0.035)		
With long tenure		-0.138***	-0.239***
		(0.053)	(0.069)
With short tenure		-0.211***	-0.150***
		(0.058)	(0.045)
Parameter tests:			
$\Pr(\beta_{M,\text{long}} = \beta_{L,\text{long}})$		0.273	0.018
$\Pr(\beta_{M,\text{long}} = \beta_{M,\text{short}})$		0.014	0.000
$\Pr(\beta_{L,\text{long}} = \beta_{L,\text{short}})$		0.106	0.113
$\Pr(\lambda_{M,\text{long}} = \lambda_{L,\text{long}})$		0.028	0.656
$\Pr(\lambda_{M,\text{long}} = \lambda_{M,\text{short}})$		0.637	0.006
$\Pr(\lambda_{L,\text{long}} = \lambda_{L,\text{short}})$		0.386	0.315
Obs.	28260	28260	28260

Notes: The dependent variable is the change in log capital-labor ratio ($\Delta \ln(K_{i,j,t}/L_{i,j,t})$). The cut off length for distinguishing between long and short tenure is equal to 12 months. Standard errors are reported in parentheses and are clustered at the firm level. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln(K_{i,j,t-1}/L_{i,j,t-1})$ is instrumented for using $\ln(K_{i,j,t-2}/L_{i,j,t-2})$. The lag length is chosen to minimize autocorrelation in the residual.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

gap should not vary with observed worker skill.

On the other hand, the productive knowledge spillover channel does predict that we may see a relationship between the measured quality of new hires and the effect of the productivity gap on the hiring firm's productivity. According to the knowledge spillover channel, workers who are skillful enough to acquire high levels of human capital (worker quality) are also likely to be able to acquire more knowledge as to how their employer operates. Alternatively more skillful workers may have greater autonomy in the reach or scope of their job within the hiring firm. Hence they may be more capable of implementing new productivity ideas in the hiring firm. Either way, more skillful workers may be able to have a large effect on productivity gains when compared to less skillful workers.

Table 2.11 reports the regression results from investigating the relationship between worker skill and the amount of productivity knowledge transferred to the hiring firm. The productivity gap and hire intensities related to hires from more and less productive firms are further divided into new variables based upon the new hire's measure of worker quality. For simplicity the new groupings are based on whether the worker's measured quality is in the top, middle, or bottom third of the economy-wide distribution of worker quality. More formally the change in the hiring firm's stock of productive knowledge is modelled as:

$$\begin{aligned}\text{Exposure}_{i,t} &= \sum_{\text{skill}\in\{\text{low,med,high}\}} \beta_{M,\text{skill}} \frac{\sum_{n\in\mathcal{N}^M_{j,t-1}} \mathbb{D}_{\text{skill}}(n)\mathbb{D}_n[\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\ &+ \sum_{\text{skill}\in\{\text{low,med,high}\}} \beta_{L,\text{skill}} \frac{\sum_{n\in\mathcal{N}^L_{i,t-1}} \mathbb{D}_{\text{skill}}(n)[\ln(A_{n,\tau(n)}) - \ln(A_{i,t-1})]}{L_{i,t-1}} \\ &+ \sum_{s\in\mathcal{S}_{i,t-1}} \lambda_s \frac{H_{i,s,t-1}}{L_{i,t-1}}, \qquad (2.11)\end{aligned}$$

where 'skill = low' denotes a worker skill in the bottom third of the distribution of worker quality, 'skill = med' denotes a worker skill in the middle third of the distribution of worker quality, and 'skill = high' denotes a worker skill in the top third of the distribution of worker quality. $\mathbb{D}_{\text{skill}}(n)$ is a dummy variable that takes on the value 1 if worker n is in the skill third of the distribution of worker quality.

For all three productivity measures in Table 2.11, the productivity gap associated with hiring low skilled workers from more productive firms has a smaller influence on the hiring firm's productivity growth than the productivity gaps associated with hiring medium and high skilled workers from more productive firms (although the difference is generally not statistically different). For both the value-added and trans-log MFP measures of productivity, the coefficient for the productivity gap associated with low skilled hires from more productive firms is around half that of the coefficient for other skill groups. So for example, if a firm with a hiring intensity from more productive firms of 10 percent was hiring low skilled workers, raising the average productivity of the firms these workers were sourced from by one percent would be associated with a 0.38 percentage point increase in labor productivity on average. While if the firm instead hired medium or high skilled workers, raising the average productivity of the firms these workers were sourced from by one percent would be associated with around a 0.7 percentage point increase in firm productivity growth.

The coefficients related to productivity gap for hires from less productive firms are similar in magnitude across all skill levels for each of the productivity measures, and are not statistically different from each another. The expected gain in productivity growth when improving the productivity of the less productive firms workers were sourced from would not differ across the various worker quality categories.

Taken at face value, the implications of these results do not line up directly with the predictions from either the knowledge spillover or unmeasured worker quality channels. In terms of the knowledge spillover channel, while we do see a larger coefficient on the productivity gap from more productive firms when comparing medium to low skilled workers, we do not see the same pattern when comparing medium to high skilled workers. In addition, this relationship between worker skill and the productivity gap effect is not isolated to labor productivity, where we have seen previous support for the knowledge spillover channel, but

Table 2.11: Worker flows by worker skill level

	Value-added		Cobb-Douglas		Trans-log	
	Baseline	**By skill Group**	**Baseline**	**By skill Group**	**Baseline**	**By skill Group**
Productivity gap, hires from (β):						
More prod. firms	0.480***		0.271***		0.354***	
	(0.098)		(0.065)		(0.068)	
Low skilled		0.380**		0.255		0.225
		(0.179)		(0.170)		(0.147)
Medium skilled		0.701**		0.348**		0.673***
		(0.273)		(0.140)		(0.164)
High skilled		0.700***		0.310**		0.466***
		(0.179)		(0.142)		(0.157)
Less prod. firms	0.153***		0.374***		0.374***	
	(0.030)		(0.054)		(0.056)	
Low skilled		0.097		0.477**		0.427**
		(0.094)		(0.204)		(0.167)
Medium skilled		0.134		0.497**		0.517***
		(0.109)		(0.232)		(0.176)
High skilled		0.159*		0.444***		0.359**
		(0.085)		(0.170)		(0.159)
Hire intensity (λ):						
More prod. firms	-0.200***		-0.012		-0.037*	
	(0.057)		(0.028)		(0.021)	
Less prod. firms	-0.117***		0.047*		0.004	
	(0.027)		(0.026)		(0.019)	
Low skilled		-0.083		0.055		-0.005
		(0.058)		(0.065)		(0.039)
Medium skilled		-0.210**		0.001		-0.040
		(0.089)		(0.062)		(0.038)
High skilled		-0.268***		0.040		-0.037
		(0.072)		(0.057)		(0.037)
Unknown skill		0.702		0.262		0.480
		(0.920)		(0.501)		(0.349)
Parameter tests:						
$\Pr(\beta_{M,low} = \beta_{L,low})$		0.241		0.496		0.457
$\Pr(\beta_{M,med} = \beta_{L,med})$		0.105		0.641		0.580
$\Pr(\beta_{M,high} = \beta_{L,high})$		0.022		0.608		0.682
$\Pr(\beta_{M,low} = \beta_{M,med} = \beta_{M,high})$		0.408		0.907		0.130
$\Pr(\beta_{L,low} = \beta_{L,med} = \beta_{L,high})$		0.909		0.985		0.834
Obs.	37269	37269	28260	28260	38037	38037

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$, where the measure of productivity differs by column. Low, medium, and high skill denotes which third of distribution of worker quality an individual is in relative to the population at the time of hiring. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

also affects the MFP results in which we have not seen previous support for a knowledge spillover channel.

In terms of the unmeasured worker quality channel, the baseline results showed support for unmeasured worker quality influencing firm MFP growth. However, the results in Table 2.11 suggest that the productivity gap for low skilled workers has a different effect from the productivity gaps of medium and high skilled workers, contradicting the predictions of the unmeasured worker quality channel.

One possible explanation for the contradictions found above is that low skilled labor is utilised differently in the production process when compared to medium and high skilled labor (e.g. only low skill workers perform manual labor jobs while higher skilled labor is used to perform other tasks). The measures of productivity so far assume that labor is a homogenous input into the production process. If low skilled labor is in effect utilised differently to medium and high skilled labor, the production function specifications used may not fully capture the distinction between low and other skilled labor inputs. This will affect the measures of productivity and hence the estimated productivity gains associated with hiring workers of different skill.

If we assume that low skilled workers are a different type of production input for the firm, the results above suggest that worker skill (the distinction between medium and high skill) does not affect the productivity gap for either MFP or labor productivity. This is consistent with the unmeasured worker quality channel. Furthermore, because the value-added and MFP results do not differ dramatically, this similarity also suggests that worker skill is not an important determinant of the productivity knowledge spillover seen in labor productivity results.

2.7.3 Robustness

The regression analysis conducted so far has only identified correlations between the hiring of new workers and the subsequent productivity growth in the hiring firm. While these correlations are consistent with the predictions of a worker quality channel and knowledge spillover, the regressions alone do not imply causality. It is possible that the causality runs in the other direction, i.e. decisions at the firm, or other outside factors are inducing productivity shocks, which in turn drive the observed hiring patterns. The analysis in this section attempts to provide some control for causality within the model, and deal with other issues of robustness of the results.[26]

Reverse Causality – Estimates of Shocks

To properly control for the direction of causality would require either knowing the productivity shocks that the firm observes when making its hiring decisions or knowing the reason for each new hire. While this level of control is not possible in the data, several techniques have been developed in the literature that attempt to identify the productivity shocks the firm observes, but that are hidden to the econometrician. For example, Levinsohn and Petrin (2003) developed a model that assumes observing changes in the firm's choice of material inputs into the production function provides information on the productivity shocks observed by the firm. Using the technique they developed, it is possible to back out estimates of the productivity shocks the firm observes before choosing labor and capital inputs, allowing us to estimate the component of MFP excluding the productivity shocks observed by the firm, thereby avoiding the reverse causality.[27]

Table 2.12 compares the regression results found using the MFP measure from the Cobb-Douglas production function and that productivity measure found using the Levinsohn and Petrin (2003) technique. The Cobb-Douglas results are used as the point of comparison here as the Levinsohn and Petrin (2003) approach uses a Cobb-Douglas production function to

26. Additional robustness checks are presented in Appendix **??**.

27. Another common approach in the literature is that developed by Olley and Pakes (1996). However, the necessary investment data is only collected in the data through the Annual Enterprise Survey (AES), which is only available for a subset of firms in the data. Therefore the approach of Levinsohn and Petrin (2003) is favored as it provides a larger sample size to work with.

estimate MFP.[28] In the Levinsohn and Petrin (2003) results, the coefficients on the productivity gap variable are slightly lower for both hires from more productive (0.211 vs 0.271) and less productive (0.213 vs 0.374) firms. While the difference in magnitudes between the coefficient on the productivity gaps for hires from more and less productive firms was not statistically significant in the Cobb-Douglas case, the two coefficients have become more similar in size after controlling for the productivity shocks observed by the firm. This brings the results in line with those based on the trans-log productivity measure. As a result, after controlling for the Levinsohn and Petrin (2003) productivity shocks, the Cobb-Douglas model estimation results provide slightly stronger support in favor of the unmeasured worker quality channel.

2.7.4 Additional instruments for $\Delta \ln A_{i,j,t-1}$

Lagged productivity is an important control in the model. Today's productivity growth as well as the firm's hiring decisions are likely to be correlated with lagged productivity. However, as mentioned in Section 2.5, including $\Delta \ln A_{i,j,t-1}$ in the regression is problematic due to the presence of large Nickell bias in the data. Therefore, in the baseline model $\ln A_{i,j,t-2}$ is used as an instrument for $\Delta \ln A_{i,j,t-1}$. However, $\ln A_{i,j,t-2}$ is not the only instrument that can be used for $\Delta \ln A_{i,j,t-1}$. Blundell and Bond (1998) developed a technique that extends the Arellano-Bond estimation to include both the lagged levels and lagged differences of productivity as suitable instrumental variables in the estimation of dynamic panel models.

Table 2.13 reports the results of estimating the model using the Blundell and Bond (1998) approach with extra instruments for the change in lagged productivity ($\Delta \ln A_{i,j,t-1}$) against the baseline regressions. For all three productivity measured considered, the coefficients in the model do not change dramatically with the inclusion of additional instruments for past

28. The Stata function 'levpet' developed by Petrin et al. (2004) is used to construct the Levinsohn-Petrin productivity measure.

Table 2.12: Effect of controlling for unobserved productivity shocks

	Cobb-Douglas	**Levinsohn-Petrin**
Productivity gap, hires from (β):		
More prod. firms	0.271***	0.211***
	(0.065)	(0.037)
Less prod. firms	0.374***	0.213***
	(0.054)	(0.056)
Hire intensity (λ):		
More prod. firms	-0.012	0.021
	(0.028)	(0.025)
Less prod. firms	0.047*	-0.010
	(0.026)	(0.036)
$\Delta Q_{i,t}$ due to (γ):		
New hires	0.105*	0.122**
	(0.062)	(0.060)
Exiters	0.103*	0.120**
	(0.062)	(0.060)
Incumbents	0.110*	0.120**
	(0.062)	(0.059)
Parameter tests:		
$\Pr(\beta_M = \beta_L)$	0.217	0.983
$\Pr(\lambda_M = \lambda_L)$	0.145	0.501
Obs.	28260	38037

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$, where the measure of productivity differs by column. The Levinsohn-Petrin measure of productivity is derived using the method developed by Levinsohn and Petrin (2003). Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

productivity. This finding suggests the results of our analysis are not sensitive to the inclusion of additional instruments.

2.8 Conclusions

The analysis carried out within this paper shows that when hiring new workers, the productivity of a worker's previous main employer is significantly correlated with the future productivity growth of the hiring firm. The strength of this correlation varies with the measure of firm productivity used, whether the firm hires from more or less productive firms, and with some characteristics of the new workers themselves.

When firm productivity is measured in terms of labor productivity (value-added per worker), the productivity gain associated with raising the average productivity of the firms that new workers are sourced from is, on average, large (small) if the hiring firm is sourcing its new workers from firms that were more (less) productive than the hiring firm. This 'premium' for hiring from more productive firms tend to be larger when the new hires are from the same industry as the hiring firm, the new hires have spent more than one year at both their previous firm and the hiring firm, and the new hires have a medium to high level of worker quality. This premium is also observed when measures of input intensity (the capital-labor ratio) is used instead of productivity.

In terms of multi-factor productivity, the productivity gain associated with improving the average productivity of firms that new workers are sourced from is independent of whether the increase in average productivity is driven by improvements to the productivity at the top or bottom ends of the distribution of source firms. Extensions to the baseline regression reveal that the magnitude of these gains are not dramatically affected by the characteristics mentioned above.

While these regression based correlations do not imply causality, it is still interesting to compare these findings to the predictions made by different models of how new hires influence firm productivity. The multi-factor productivity results are consistent with predictions of a

Table 2.13: Effect of additional instruments for lagged productivity

	Value-added		**Cobb-Douglas**		**Trans-log**	
	Baseline	**Blundell-Bond**	**Baseline**	**Blundell-Bond**	**Baseline**	**Blundell-Bond**
Productivity gap, hires from (β):						
More prod. firms	0.480***	0.515***	0.271***	0.257**	0.354***	0.250***
	(0.098)	(0.124)	(0.065)	(0.112)	(0.068)	(0.093)
Less prod. firms	0.153***	0.172***	0.374***	0.363***	0.374***	0.296***
	(0.030)	(0.043)	(0.054)	(0.086)	(0.056)	(0.075)
Hire intensity (λ):						
More prod. firms	-0.200***	-0.193***	-0.012	-0.018	-0.037*	-0.027
	(0.057)	(0.063)	(0.028)	(0.030)	(0.021)	(0.021)
Less prod. firms	-0.117***	-0.108***	0.047*	0.044	0.004	0.007
	(0.027)	(0.027)	(0.026)	(0.027)	(0.019)	(0.018)
Obs.	37269	49920	28260	38049	38037	50874

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$, where the measure of productivity differs by column. The Blundell-Bond columns report regression results using lags of productivity ($\ln(A_{i,t-x})$) and change in productivity $\ln(A_{i,t-x})$ as instruments for past productivity. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

worker quality/screening channel where more productive firms are either better at screening good quality workers or provide them with better training (creating positive assortative matching). Because the productivity of a worker's previous employer acts as a signal of worker quality, such a model would predict that the coefficients on the productivity gaps for hires from both more and less productive firms should be positive, and equal. In terms of the labor productivity results, the premium in the coefficient on the productivity gap for hires from more productive firms seen in the value-added productivity measure is consistent with the knowledge spillover channel. This channel predicts that workers from more productive firms are able to transmit new, better productivity ideas to the hiring firm. The fact that we see this relationship in the labor productivity measure, and the capital-labor ratio, but not the MFP data, would suggest the knowledge spillover is related to knowledge about production technology (more capital intensive production methods) not MFP (how to utilize the firm's current inputs more efficiently).

With the data that is available it is not possible to say definitively that it is the knowledge of new workers driving productivity growth in the hiring firms. The regressions results appear to be robust to further disaggregation of the productivity gap as well as various attempts to control for the direction of causality. This suggests that the empirical findings are at least consistent with hiring firms benefiting from the knowledge of new hires through both the worker quality and a knowledge spillover channels.

APPENDIX A

FURTHER FIRM-LEVEL SUMMARY STATISTICS

Examining the relationship between new hires and productivity growth at the hiring firm is only a worthwhile exercise if: (i) there is cross-sectional variation in firm productivity so that firm have exposure to different productive ideas when hiring, and (ii) there is dynamic variation in firm productivity so that we may try to relate changes in productivity to changes in hiring rates of new workers. This appendix explores these issues and provides more detail on the productivity of firms within New Zealand.

A.1 Distribution of Firm Productivity

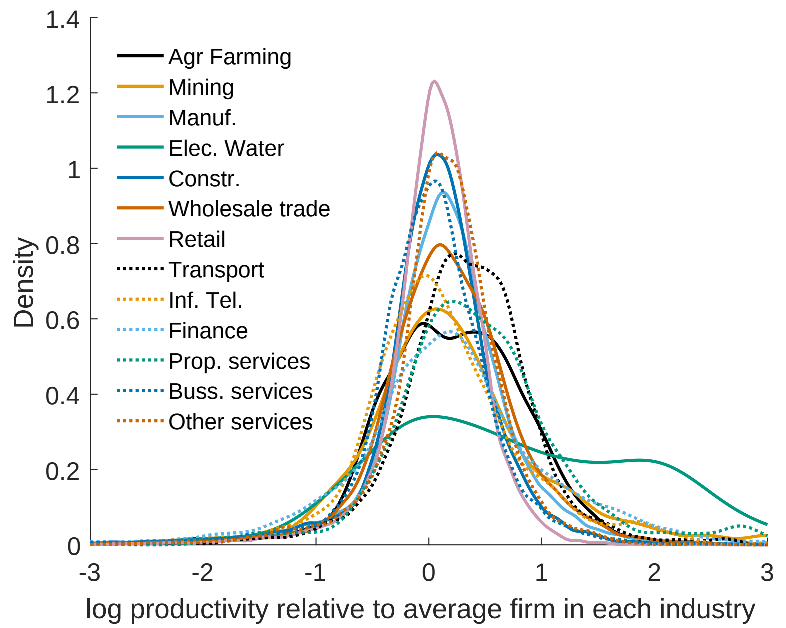
Figure A.1 plots the kernel density estimates of the various productivity distributions, aggregated up to the 1-digit industry group classifications for the firms in the sub-sample (firms with an average labor force size of at least 10 full time employees over the year), and averaged over the entire sample period 2001-2012.

The measures of MFP in most industries is distributed fairly symmetrically around the industry averages, with the Trans-log (the most flexible production function specification) showing the greatest symmetry. The notable exception to this symmetry is the Finance industry in which the productivity distribution is skewed to the right in all the productivity measures.

In terms of value-added per worker (labor productivity), subplot A.1a shows, unsurprisingly, that firms in capital intensive industries such as Mining, Electrical and Water supply, and Financing tend to have higher levels of labor productivity than more labor intensive firms such as Retail, and the various service industries. The log-productivity differences between these industries indicates that there is significant differences between the levels of value-added per worker across the different industries.

Figure A.1: Kernel density estimates of the productivity distributions

(a) Value-added per worker



(b) Cobb-Douglas

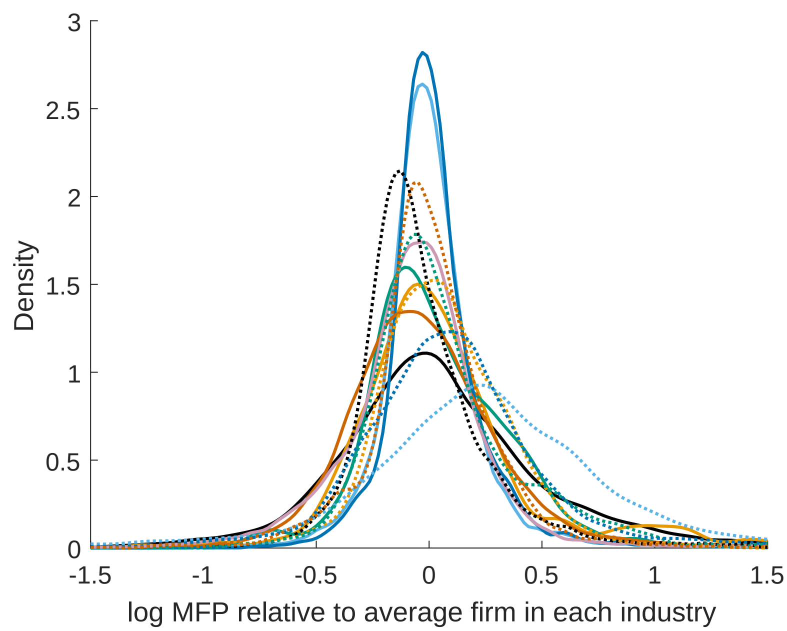
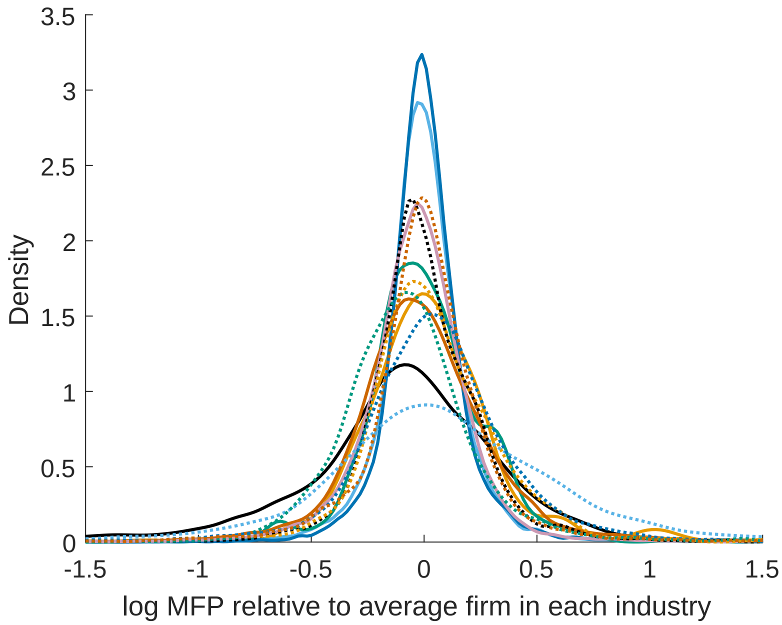


Figure A.1: Kernel density estimates of the productivity distributions (continued)

(c) Trans-log



Notes: Each subplot shows the kernel density estimate of productivity in each of the 13 1-digit industry groups for all the firms in the sample (those with an average annual employment of more than 10 full time equivalent workers) aggregated over the entire sample period. All measures of firm-level productivity are computed at the 4-digit industry level before aggregation. For the measures of MFP (Cobb-Douglas, Fixed effects, and Trans-log), firm-level productivity is computed relative to the average in the industry-year using all firms (even those with FTE<10). For the measure of value-added per worker, results from each year are converted to real values.

A.2 Productivity Dynamics

Table A.1 presents a summary of the productivity transition dynamics within the sample. For each productivity decile in a given year (row), table A.1 shows the fraction of firms within that productivity decile that where in each source (column) during the previous year. There are 12 possible sources for firms, 10 productivity deciles, and two reasons for being out of scope, either missing productivity data during the previous year, or being too small (less than 10 full time workers on average over the year).

Table A.1: Firm productivity transition matrices

(a) Value-added per worker

Firm's current	**Firm's previous productivity decile**										**Missing**	**L<10**
prod. decile	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**prod data**	
1	0.32	0.13	0.05	0.04	0.03	0.02	0.02	0.01	0.01	0.01	0.27	0.1
2	0.11	0.31	0.14	0.05	0.02	0.02	0.01	0.01	0.01	0	0.22	0.1
3	0.05	0.13	0.26	0.15	0.06	0.03	0.02	0.01	0	0	0.2	0.09
4	0.04	0.05	0.14	0.22	0.15	0.07	0.03	0.02	0.01	0	0.2	0.09
5	0.03	0.02	0.05	0.14	0.23	0.14	0.06	0.03	0.01	0.01	0.18	0.09
6	0.02	0.02	0.03	0.06	0.14	0.22	0.16	0.06	0.02	0.01	0.18	0.07
7	0.01	0.01	0.01	0.03	0.06	0.14	0.24	0.16	0.05	0.01	0.19	0.07
8	0.01	0.01	0.01	0.01	0.03	0.06	0.15	0.27	0.17	0.03	0.18	0.07
9	0.01	0	0	0.01	0.01	0.02	0.05	0.15	0.35	0.13	0.19	0.07
10	0.01	0	0	0	0	0.01	0.01	0.03	0.12	0.57	0.19	0.06

(b) Cobb-Douglas

Firm's current	**Firm's previous productivity decile**										**Missing**	**L<10**
prod. decile	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**prod data**	
1	0.33	0.12	0.06	0.03	0.03	0.02	0.02	0.02	0.01	0.02	0.25	0.09
2	0.12	0.25	0.15	0.07	0.04	0.03	0.02	0.02	0.01	0.01	0.21	0.08
3	0.05	0.13	0.2	0.15	0.08	0.04	0.03	0.02	0.01	0.01	0.2	0.08
4	0.04	0.07	0.13	0.18	0.13	0.08	0.04	0.03	0.02	0.01	0.2	0.08
5	0.03	0.04	0.07	0.13	0.18	0.14	0.08	0.05	0.02	0.01	0.18	0.08
6	0.02	0.03	0.04	0.07	0.13	0.17	0.14	0.07	0.03	0.02	0.19	0.09
7	0.02	0.02	0.02	0.04	0.07	0.14	0.18	0.14	0.07	0.02	0.19	0.08
8	0.01	0.01	0.02	0.02	0.04	0.07	0.13	0.22	0.16	0.04	0.18	0.09
9	0.01	0.01	0.01	0.01	0.02	0.03	0.06	0.14	0.28	0.14	0.18	0.09
10	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.04	0.13	0.44	0.2	0.09

The transition matrices in table A.1 show that there is some persistence in the firm's productivity ranking. Depending upon the productivity measure and productivity decile, a

Table A.1: Firm productivity transition matrices (continued)

(c) Trans-log

Firm's current prod. decile	**Firm's previous productivity decile 1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Missing prod data**	**L<10**
1	0.32	0.13	0.06	0.04	0.03	0.02	0.02	0.02	0.01	0.01	0.26	0.09
2	0.12	0.24	0.15	0.07	0.05	0.03	0.02	0.02	0.01	0.01	0.21	0.08
3	0.06	0.14	0.19	0.13	0.08	0.05	0.03	0.02	0.01	0.01	0.2	0.08
4	0.04	0.07	0.13	0.18	0.13	0.08	0.05	0.03	0.02	0.01	0.19	0.08
5	0.03	0.04	0.08	0.13	0.17	0.13	0.08	0.05	0.03	0.01	0.19	0.07
6	0.02	0.03	0.05	0.08	0.13	0.16	0.13	0.08	0.04	0.02	0.2	0.08
7	0.02	0.02	0.02	0.04	0.07	0.13	0.18	0.15	0.07	0.03	0.18	0.09
8	0.01	0.01	0.02	0.02	0.04	0.07	0.14	0.21	0.16	0.05	0.18	0.09
9	0.01	0.01	0.01	0.01	0.02	0.04	0.06	0.14	0.28	0.14	0.19	0.08
10	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.05	0.14	0.43	0.2	0.09

Notes: Each cell shows the fraction firms in the productivity decile for the current year (row) that were in each productivity decile, or out of scope in the previous year (column). For example, cell (1,1) refers to the fraction of firms in productivity decile one, that were also in productivity decile one last year. And cell (1,2) refers to the fraction of firms in productivity decile one that were in productivity decile two last year. Cells are shaded based upon the fraction of firms that were in that previous source last year, with darker shades corresponding to a higher fraction. Deciles correspond to the firm's productivity ranking within each year, with decile 10 referring to the most productive firms.

firm has around a 20 to 50 percent chance of been in the same productivity decile in the previous year. If firms do transition between productivity deciles, they tend not to make large jumps between very different deciles. And this pattern holds for all of the productivity deciles considered.

A.3 Distribution of Firm Size (Zipf's Law)

For many countries around the world, the distribution of firm sizes (measured in labor units) in the economy has been shown to be well approximated by a Pareto distribution where the tail parameter is close to unity. This implies that the share of firms whose size is above a given value is inversely proportional to that value. More formally, The share of firms with a size larger than s is given by

$$\Pr[S \geq s] = \left(\frac{\lambda}{s}\right)^{\alpha} \tag{A.1}$$

where s is the size of the firm (measured in labor units), λ is a scale parameter, and α is the tail parameter. Zipf's law implies $\alpha = 1$.

To test whether the Zipf's law holds for the distribution of firm size in New Zealand, the log of the firm's size percentile is regressed on the log of the firm size and a constant. The coefficient on the log of the firm size corresponds to $-\alpha$. The results of this regression are presented in Table A.2. Because there are a significant number of very small sized firms (and firms without employees) in the economy, the regression is run on several different sub samples where firm's below a certain minimum size have been dropped.

Table A.2: Estimates of Zipf's law for the distribution of firm size

	Minimum firm size (FTE =)			
	1	**2**	**5**	**10**
log(FTE) ($-\alpha$)	-0.963***	-1.044***	-1.089***	-1.084***
	(0.000)	(0.000)	(0.000)	(0.000)
constant	-0.128***	0.803***	1.761***	2.486***
	(0.000)	(0.000)	(0.000)	(0.000)
Pr($\alpha = 1$)	0	0	0	0
R^2	0.993	0.997	0.999	0.997
Obs.	1284588	608814	265680	126288

Notes: Dependent variable is the log of the firm size percentile. Each column represents a separate regression for sub-samples using different minimum firm sizes, where firm size is measured in average Full Time Equivalent workers (FTE) over the year. A coefficient of $\alpha = -1$ is consistent with Zipf's law. Standard errors are in parentheses.
* $p < 0.1$, ** $p < 0.02$, *** $p < 0.01$

The point estimates of $-\alpha$ for all the sub-sample regressions are close in magnitude to negative one, the value implied by Zipf's law, although given the large number of observations in the sample their values are statistically different from minus one at the one percent level of significance. A large number of firms in the data set employ one or fewer full time employees on average through the year (e.g. some sole proprietors). Increasing the minimum firm size for the regression from one to two FTEs halves the number of observations, and has a noticeable impact on the point estimate of α, increasing it form 0.963 to 1.044. However, the estimate of α is relatively robust to further increases in the minimum firm size used for the sample and remains marginally greater than unity. As a result, the distribution of firm sizes in the New Zealand economy is consistent with Zipf's law.

APPENDIX B

ADDITIONAL REGRESSION SPECIFICATIONS

B.1 Alternative Measures of Worker Quality

The measure of worker quality used so far is based solely on the characteristics of the worker that are applicable to all firms (the worker fixed effect and other demographic type variables). It does not account for how productive the worker is in a specific job. As a result, the measure of worker quality could be misleading for the measure of quality applicable to the specific hiring firm.

To address this concern, the baseline model is re-estimated using two alternative measures of worker quality. The first alternative is to only use the worker fixed effect from the wage regression (dropping the characteristics such as age and gender). The second alternative is to construct a measure of worker quality by subtracting the firm fixed effect from the worker's wage (effectively adding the regression residual back into the baseline measure of quality). This second measure of worker quality should capture any worker-firm match quality at the hiring firm. Our main interest lies in seeing if the estimates of the coefficients for the productivity gaps and hire intensities vary with the changes in the choice of worker quality measure. If they do vary significantly, it would suggest that the support found so far for the knowledge spillover and unmeasured worker quality channel could be related to the lack of control within the model for the worker-firm match quality. The results of the regressions using the alternate measures of worker quality are presented in table B.1.

The regression results in table B.1 show that the coefficients related to the productivity gaps and hire intensities are not significantly affected when the measure of worker quality is changed to either of the alternatives. This set of results suggests that the productivity gap and hiring intensities are not proxying for worker-firm match quality that can be measured through the observation of the worker's wage.[1]

1. Technically it is still possible that the firms may have significant bargaining power in the wage nego-

Table B.1: Alternative measures of worker quality

	Value-added			Cobb-Douglas			Trans-log		
Measure of worker quality:	**Baseline**	**WFE**	**Wages-FFE**	**Baseline**	**WFE**	**Wages-FFE**	**Baseline**	**WFE**	**Wages-FFE**
Productivity gap, hires from (β):									
More prod. firms	0.480***	0.485***	0.482***	0.271***	0.272***	0.271***	0.354***	0.356***	0.353***
	(0.098)	(0.099)	(0.098)	(0.065)	(0.065)	(0.064)	(0.068)	(0.069)	(0.068)
Less prod. firms	0.153***	0.156***	0.151***	0.374***	0.375***	0.375***	0.374***	0.380***	0.373***
	(0.030)	(0.031)	(0.030)	(0.054)	(0.054)	(0.054)	(0.056)	(0.056)	(0.056)
Hire intensity (λ):									
More prod. firms	-0.200***	-0.180***	-0.201***	-0.012	-0.008	-0.011	-0.037*	-0.028	-0.037*
	(0.057)	(0.058)	(0.057)	(0.028)	(0.027)	(0.028)	(0.021)	(0.021)	(0.021)
Less prod. firms	-0.117***	-0.106***	-0.110***	0.047*	0.050*	0.048*	0.004	0.011	0.005
	(0.027)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.019)	(0.018)	(0.019)
Parameter tests:									
$\Pr(\beta_M = \beta_L)$	0.001	0.001	0.001	0.217	0.221	0.211	0.808	0.785	0.811
$\Pr(\lambda_M = \lambda_L)$	0.237	0.293	0.196	0.145	0.147	0.137	0.174	0.208	0.164
Obs.	37269	37269	37269	28260	28260	28260	38037	38037	38037

Notes: Dependent variable is the regressions is the change in log productivity ($\Delta \ln A_{i,j,t}$). Each column denotes a different measure of worker quality ($Q_{i,t}$). WFE denotes worker quality measured by the Worker Fixed Effect. Wages-FFE denotes worker quality measured as the log of wages less the Firm Fixed Effect. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

B.2 Average Productivity Differences Between Industries

A possible reason of why the estimation based on value-added per worker data supports the knowledge spillover channel while the results based on MFP measures do not is that between-industry productivity differences are important. As discussed in section 2.5.3 of the paper, constructing the productivity gaps using MFP measures fails to account for any differences in the average productivity between industries. Therefore, the productivity gap measures could be misleading as measures of the true productivity differences.

To investigate whether the MFP estimates are biased as a result of using productivity gap measures that do not account for the between-industry productivity differences, the baseline model is re-estimated using value-added per worker data that is demeaned by the industry-year average. By demeaning in this manner, the between-industry productivity differences are removed from productivity gaps in the same manner as they are for the MFP measures of productivity. By comparing the results from the demeaned value-added labor productivity measure to the original value-added labor productivity results, we should be able to see if between-industry productivity differences significantly affect the estimated. The results of this comparison are presented in table B.2.

All of the key coefficients in the two columns of table B.2 are similar in magnitude and direction. This suggests that the difference in productivity gap coefficients between the value-added per worker and the MFP measures of firm productivity are not being driven by the fact that productivity gaps based on MFP measures fail to account for between-industry differences in average productivity. Instead the differences are likely the result of how MFP measures treat other inputs in the production process.

---

tiations that they are able to fully capture the benefit of worker skill without paying the workers a higher wage. However, with the regulatory environment in New Zealand, it is unlikely firms have this much power.

Table B.2: Baseline model estimated using demeaned value-added data

	Value-added	**VA pw demeaned**
Productivity gap, hires from (β):		
More prod. firms	0.480***	0.488***
	(0.098)	(0.108)
Less prod. firms	0.153***	0.139***
	(0.030)	(0.028)
Hire intensity (λ):		
More prod. firms	-0.200***	-0.210***
	(0.057)	(0.061)
Less prod. firms	-0.117***	-0.116***
	(0.027)	(0.028)
$\Delta Q_{i,t}$ due to (γ):		
New hires	0.479***	0.468***
	(0.075)	(0.075)
Exiters	0.468***	0.457***
	(0.075)	(0.075)
Incumbents	0.494***	0.483***
	(0.075)	(0.074)
Parameter tests:		
$\Pr(\beta_M = \beta_L)$	0.001	0.001
$\Pr(\lambda_M = \lambda_L)$	0.237	0.221
$\Pr(\gamma_{new} = \gamma_{incmb})$	0.000	0.000
Obs.	37269	37269

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$, where the measure of productivity differs by column. Demeaned value-added per worker is demeaned using industry-year averages. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

B.3 Stoyanov and Zubanov (2012) Estimation Specification

While the baseline model used in this paper has similarities to the model used by Stoyanov and Zubanov (2012), an important difference is that (2.4) relates the productivity gap to the change in productivity at the hiring firm, while the model of Stoyanov and Zubanov (2012) relates the productivity gap to the level of productivity at the hiring firm. The results from the analysis of this paper are also different to the results of the analysis by Stoyanov and Zubanov (2012) for Danish data. This paper finds support for the unmeasured worker quality channel in both labor productivity and MFP measures, and some support for the knowledge spillover channel only in the labor productivity data. Stoyanov and Zubanov (2012) on the other hand finds support for only the knowledge spillover channel, both in labor productivity and MFP data.

To eliminate the possibility that the differences in our findings and those of Stoyanov and Zubanov (2012) are the result of the choice of modelling approach, we transform the baseline model given in (2.4) to a form that is closer to the structure the model used by Stoyanov and Zubanov (2012), and then re-estimate the model using the new form.

Substituting the identity $\Delta \ln A_{i,j,\tau(n)} = \ln A_{i,j,\tau(n)} - \ln A_{i,j,\tau(n)-1}$ into (2.4) and rearranging yields the following expression which, like the model of Stoyanov and Zubanov (2012), relates the level of MFP to the change in firm knowledge:

$$\ln A_{i,j,t} = \text{Exoposure}_{i,t} + \gamma \Delta Q_{i,t} + \delta \Delta ExTurn_{i,t} + \sum_{l=1}^{L} \beta_{A,l} \ln A_{i,j,t-l} + \Delta\theta_{j,t} + \varepsilon_{i,t} \quad \text{(B.1)}$$

Following the approach of Stoyanov and Zubanov (2012), the dynamic panel model relationship above is estimated using a first-difference approach. Like for the baseline model in the main part of the paper, the lagged level of productivity in period $t-2$ is used to instrument $\Delta \ln A_{i,j,t-1}$ as a control for potential Nickell bias in the regression. The results of the regression using all three firm productivity measures are presented in table B.3.

The coefficients in table B.3 for the Stoyanov and Zubanov (2012) like model differ in

Table B.3: Regressions using a functional form similar to Stoyanov and Zubanov (2012)

	Value-added		**Cobb-Douglas**		**Trans-log**	
	Baseline	**S-Z like**	**Baseline**	**S-Z like**	**Baseline**	**S-Z like**
Prod. gap, hires from (β):						
More prod. firms	0.480***	0.909***	0.271***	0.459***	0.354***	0.654***
	(0.098)	(0.092)	(0.065)	(0.070)	(0.068)	(0.090)
Less prod. firms	0.153***	0.147***	0.374***	0.673***	0.374***	0.397***
	(0.030)	(0.026)	(0.054)	(0.077)	(0.056)	(0.065)
Hire intensity (λ):						
More prod. firms	-0.200***	-0.319***	-0.012	0.042	-0.037*	-0.053*
	(0.057)	(0.045)	(0.028)	(0.026)	(0.021)	(0.025)
Less prod. firms	-0.117***	-0.190***	0.047*	0.073**	0.004	-0.049*
	(0.027)	(0.024)	(0.026)	(0.026)	(0.019)	(0.019)
Parameter tests:						
$\Pr(\beta_M = \beta_L)$	0.001	0.000	0.217	0.026	0.808	0.014
$\Pr(\lambda_M = \lambda_L)$	0.237	0.009	0.145	0.382	0.174	0.900
Obs.	37269	49905	28260	50859	38037	50859

Notes: Dependent variable is the regressions is the change in log productivity ($\Delta \ln A_{i,j,t}$). "S-Z like" refers to the model estimated by first differencing equations 2.8 and B.1, based on the model used by Stoyanov and Zubanov (2012). Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

magnitude from those of the baseline model presented earlier in the paper. However, the relative size and direction of the parameters are broadly similar.[2] Therefore, the findings in this paper are robust to the change in model specification considered by Stoyanov and Zubanov (2012). And the modelling choice approach does not seem to be the driver of the differences in results from our analysis and that of Stoyanov and Zubanov (2012).

B.4 Controlling for Worker Quality

Stockinger and Wolf (2016) found that for German firms, hiring from more productive firms was not correlated with productivity gains at the hiring firm, but hiring from less productive firms was. Their analysis reveals that there is a strong selection effect occurring, where the workers who move from more to less productive firms tend to be the workers who are among the lowest paid in their firm, and the workers who move from less to more productive firms tend to be the from the highest paid in the firm. To explore whether such a selection effect is influencing the estimates of the productivity gap and hire intensity coefficients in this paper's analysis, the approach of Stockinger and Wolf (2016) is adapted to the baseline model.

Since we are unable to observe the reason for the hiring firm's recruitment choices, it is possible that the types of workers the firms recruit are correlated with the productivity level of the worker's previous firm (e.g. hire only managers from productive firms, and production line workers from less productive firms). As a result, the coefficient on the productivity gap variables will be biased as it is measuring two factors, the productivity knowledge from new workers, and the job-type of these new workers.

As an attempt to control for this, two different measures of the worker's rank within the sending firm are used. The average ranking of new workers' pay, and the average ranking of new workers' previous human capital. Recall that the human capital measure is constructed from a log-wage regression so both rankings used as controls are correlated with one another,

2. The most significant difference is in the last column where the coefficient on the productivity gap from hires from more productive firms is now larger than the coefficient on the productivity gap from hires from less productive firms.

so one should expect that the results should not be too dis-similar. Table B.4 presents the results of estimating the baseline model with these additional controls.

The results show that the key coefficients in the regression are unaffected by the inclusion of the additional controls. Therefore, it does not appear that the knowledge gap variables is proxying for the (measurable) quality of workers within their previous firm. So unlike the German data, there does not seem to be a significant selection effect in the types of workers hired from more and less productive firms within the New Zealand data.

Table B.4: Additional controls for the types on new workers hired

	Value-added			**Cobb-Douglas**			**Trans-log**		
	Baseline	**Additional controls:**		**Baseline**	**Additional controls:**		**Baseline**	**Additional controls:**	
		real earn. ranking	**quality ranking**		**real earn. ranking**	**quality ranking**		**real earn. ranking**	**quality ranking**
Productivity gap, hires from (β):									
More prod. firms	0.480***	0.478***	0.530***	0.271***	0.269***	0.199***	0.354***	0.352***	0.266***
	(0.098)	(0.098)	(0.082)	(0.065)	(0.064)	(0.060)	(0.068)	(0.068)	(0.078)
Less prod. firms	0.153***	0.144***	0.131**	0.374***	0.374***	0.448***	0.374***	0.370***	0.324***
	(0.030)	(0.030)	(0.041)	(0.054)	(0.054)	(0.066)	(0.056)	(0.055)	(0.056)
Earnings pctile rank within:									
More prod. firms		0.043***			0.005			0.010**	
		(0.007)			(0.006)			(0.004)	
Less prod. firms		-0.034***			-0.015*			-0.010**	
		(0.009)			(0.007)			(0.005)	
Worker qual. pctile rank within:									
More prod. firms			-0.125***			0.003			0.001
			(0.019)			(0.013)			(0.009)
Less prod. firms			-0.015			-0.040***			-0.008
			(0.020)			(0.015)			(0.012)
Parameter tests:									
$\Pr(\beta_M = \beta_L)$	0.001	0.001	0.000	0.217	0.206	0.004	0.808	0.834	0.533
$\Pr(\lambda_M = \lambda_L)$	0.237	0.131	0.001	0.145	0.107	0.047	0.174	0.116	0.537
Obs.	37269	37269	23199	28260	28260	18423	38037	38037	23877

Notes: The dependent variable is $\Delta \ln A_{i,j,t}$.The average rank at the previous firm is measured as the average percentile ranking of the worker's who leave to be hired by the hiring firm. Standard errors are reported in parentheses. Each regression includes industry-year fixed effects, lagged productivity changes, hiring intensities from other sources, and excess turnover as additional regressors. The regressor $\Delta \ln A_{i,j,t-1}$ is instrumented for using $\ln A_{i,j,t-2}$. Productivity lag length is chosen to minimize autocorrelation in the residual. Productivity gaps are constructed using the subset of new hires from other private for profit firms for which productivity can be observed. The regression standard errors are clustered at the firm level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

REFERENCES

John M. Abowd, Francis Kramarz, and David N. Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

John M. Abowd, Francis Kramarz, Paul Lengermann, and Sébastien Pérez-Duarte. Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Matching Model for France and the United States. 2004.

Daron Acemoglu and William B. Hawkins. Search With Multi-Worker Firms. *Theoretical Economics*, 9(3):583–628, September 2014. ISSN 19336837. doi: 10.3982/TE1061.

Gary S. Becker. A Theory of Marriage: Part I. *The Journal of Political Economy*, 81(4): 813–846, 1973.

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

Francisco J. Buera and Robert E. Lucas, Jr. Idea Flows and Economic Growth. February 2018.

Kunal Dasgupta. Learning and Knowledge Diffusion in a Global Economy. September 2011.

Richard Fabling. Keeping it Together: Tracking Firms on New Zealand's Longitudinal Business Database. *Motu Working Paper*, 11-01, March 2011.

Richard Fabling and David C. Maré. Addressing the Absence of Hours Information in Linked Employer-Employee Data. *Motu Working Paper*, 15-17, October 2015a.

Richard Fabling and David C. Maré. Production Function Estimation Using New Zealand's Longitudinal Business Database. *Motu Working Paper*, 15-15, September 2015b.

Richard Fabling and Lynda Sanderson. A Rough Guide to New Zealand's Longitudinal Business Database (2nd edition). *Motu Working Paper*, 16-03, 2016.

Andrea Fosfuri, Massimo Motta, and Thomas Ronde. Foreign Direct Investment and Spillovers Through Workers' Mobility. January 1998.

Amy Jocelyn Glass and Kamal Saggi. Multinational Firms and Technology Transfer. *Scandinavian Journal of Economics*, 104(4):495–513, 2002.

Zvi Griliches. The Search for R&D Spillovers. *The Scandinavian Journal of Economics*, 94: S29, 1992. ISSN 03470520. doi: 10.2307/3440244.

Dean Robert Hyslop and David C Maré. Job Mobility and Wage Dynamics. *Global COE Hi-Stat Discussion Paper Series*, 107, December 2009.

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

Robert E. Lucas, Jr. On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1):3–42, 1988.

Robert E. Lucas, Jr. and Benjamin Moll. Knowledge Growth and the Allocation of Time. *Journal of Political Economy*, 122(1):1–51, 2014.

Erzo G. J. Luttmer. Eventually, Noise and Imitation Implies Balanced Growth. *Federal Reserve Bank of Minneapolis Working Paper*, 699, August 2012.

Erzo G. J. Luttmer. Four Models of Knowledge Diffusion and Growth. *Federal Reserve Bank of Minneapolis Working Paper*, 724, May 2015.

Keith McLeod, Richard Fabling, and David C. Maré. Hiring New Ideas: International Migration and Firm Innovation in New Zealand. *Motu Working Paper*, 14, 2014.

Alexander Monge-Naranjo. Foreign Firms and the Diffusion of Knowledge. *Journal of International Economics*, 87(2):323–336, 2012.

Dale T. Mortensen. Wage Dispersion in the Search and Matching Model. *American Economic Review - Papers and Proceedings*, 100(2):338–342, May 2010. ISSN 0002-8282. doi: 10.1257/aer.100.2.338.

Stephen Nickell. Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417, November 1981. ISSN 00129682. doi: 10.2307/1911408.

Emily Nix. Learning Spillovers in the Firm. November 2015.

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

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

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

Amil Petrin, Brian P. Poi, and James Levinsohn. Production Function Estimation in Stata Using Inputs to Control for Unobservables. *The Stata Journal*, 4(2):113–123, 2004.

Paul M. Romer. Increasing Returns and Long-Run Growth. *The Journal of Political Economy*, 94(5):1002–1037, October 1986.

Michel Serafinelli. Good Firms, Worker Flows and Local Productivity. 2015.

Isabelle Sin, Richard Fabling, Adam B Jaffe, David C. Maré, and Lynda Sanderson. Exporting, Innovation and the Role of Immigrants. *Motu Working Paper*, 14-15, December 2014.

Bastian Stockinger and Katja Wolf. The Productivity Effects of Woker Mobility Between Hetergoenous Firms. *IAB Discussion Paper*, 7, 2016.

Lars A. Stole and Jeffrey Zwiebel. Intra-Firm Bargaining Under Non-Binding Contracts. *Review of Economic Studies*, 63(3):375–410, 1996.

Andrey Stoyanov and Nikolay Zubanov. Productivity Spillovers Across Firms Through Worker Mobility. *American Economic Journal: Applied Economics*, 4(2):168–198, April 2012. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.4.2.168.