DISCUSSION PAPER SERIES

No. 9850

PRODUCTIVITY SPILLOVERS THROUGH LABOR MOBILITY

Tom-Reiel Heggedal, Espen R Moen and Edgar Preugschat

LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9850.php**

ISSN 0265-8003

PRODUCTIVITY SPILLOVERS THROUGH LABOR MOBILITY

Tom-Reiel Heggedal, Norwegian Business School
Espen R Moen, Norwegian Business School and CEPR
Edgar Preugschat, University of Konstanz

Discussion Paper No. 9850
March 2014

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9850

March 2014

ABSTRACT

Productivity Spillovers Through Labor Mobility*

Do firms have the right incentives to innovate in the presence of productivity spillovers? This paper proposes an explicit model of spillovers through labor flows in a framework with search frictions. Firms can choose to innovate or to imitate by hiring a worker from a firm that has already innovated. We show that if innovation firms can commit to long-term wage contracts with their workers, productivity spillovers are fully internalized. If firms cannot commit to long-term wage contracts, there is too little innovation and too much imitation in equilibrium. Our model is tractable and allows us to analyze welfare effects of various policies in the limited commitment case. We find that subsidizing innovation and taxing imitation improves welfare.Moreover, allowing innovation firms to charge quit fees or rent out workers to imitation firms also improves welfare. By contrast, non-pecuniary measures like covenants not to compete, interpreted as destruction of matches between imitation firms and workers from innovation firms, always reduce welfare.

JEL Classification: 031, J63, J68 and O38
Keywords: efficiency, imitation, innovation, productivity, search frictions, spillovers and worker flows

Tom-Reiel Heggedal
Department of Economics
Norwegian Business School
0442 Oslo
NORWAY

Email: tom-reiel.heggedal@bi.no

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=178082

Espen R Moen
Department of Economics
Norwegian School of Economics
0442 Oslo
NORWAY

Email: espen.r.moen@bi.no

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=149748

Edgar Preugschat
Department of Economics
University of Konstanz
Box 145
78457 Konstanz
GERMANY

Email: edgar.preugschat@uni-konstanz.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=178083

*The authors would like to thank Javier Fernandez-Blanco, Dale T. Mortensen, participants at the department seminars at the Durham University Business School, the Frisch Centre in Oslo, the Norwegian Business School in Oslo, the Norwegian School of Economics in Bergen, and the University of Oslo, and participants at the Oslo-Århus Workshop in 2011, the Search and Matching Workshop in Bristol 2011, the EEA-ESEM conference in Malaga 2012, the Mainz Workshop in Labour Economics 2012, the SED in Cyprus 2012, and at the SEAT conference in Paris 2013, for helpful comments and fruitful discussions. Thanks to the Norwegian Research Council for financial support.

Submitted 24 January 2014

1 Introduction

Productivity spillovers associated with R&D play a major role in the modern literature on economic growth.[1] Due to such productivity spillovers, the argument goes, R&D gives rise to positive externalities on other firms, which in turn may call for policies that spur innovation. An important channel for these spillovers, already noted by Arrow (1962), is labor mobility.[2] If a worker moves from a technologically advanced firm to one that is less so, she may bring valuable knowledge with her. Hence worker flows create information flows.

However, in the literature on economic growth, information flows through worker flows are not modeled. Hence it is a priori not clear under what circumstances information flows will create positive externalities from innovations, and if they do, what sorts of policy instruments would be appropriate. To tackle this question we explicitly model worker turnover within a frictional labor market, where both workers with knowledge and firms in the need of knowledge search to form a match.

More specifically, in our model competitive firms may either innovate in the first period or wait to imitate in the second period by hiring a worker from a firm that has innovated. An innovation firm that loses a worker still possesses the required knowledge, and can therefore hire a new worker that does not. All hiring is subject to costly search frictions. That is, firms pay to post a vacancy specifying a wage contract, and workers direct their search to offers.

From a social planner's perspective, there is a trade-off between innovation costs on the one hand and search- and waiting costs on the other. If a large fraction of the firms innovate, aggregate innovation costs are high. On the other hand, innovations come in more quickly and the planner economizes on search costs, as less job-to-job transitions are necessary in order to disseminate the knowledge to imitating firms. The optimal trade-off features both innovation and imitation. We show that the welfare properties of the equilibrium allocation depend on what restrictions we impose on the wage contracts offered by innovating firms. If an innovating firm can commit to long-term wage contracts, it will give the employee the full match surplus of the second period. This will induce the employee to search in a way that maximizes this surplus, which the firm in

[1]See Romer (1990), Grossman and Helpman (1993) and Aghion and Howitt (1992). For a survey of the literature on growth and spillovers see Jones (2005).

[2]For the empirical evidence see our literature discussion below.

turn extracts through a relatively low period-1 wage. As a result a firm that innovates pockets the full social value of its innovation, and the decentralized equilibrium realizes the socially optimal allocation. If instead firms cannot commit to future wages, they offer lower wages at the on-the-job search stage in the second period. Imitation by hiring workers becomes cheaper, implying too much imitation and too little innovation in equilibrium compared with the social optimal levels. Without search frictions, the allocation is still efficient. Hence, the combination of search frictions and limited commitment creates inefficiency.

In the paper we develop a new methodology for welfare analysis of search models in the absence of long-term wage contracts.[3] We find that a subsidy to innovators, together with a tax on imitation, can implement the efficient allocation. If only a subsidy to innovating firms or only a tax on imitating firms is used, welfare can be increased, but not all distortions can be corrected. This may seem trivial, however, the results crucially hinge on general equilibrium effects between labor markets.

Importantly, we also study the welfare implications of (ex-post) firm-level measures aimed at reducing excessive turnover. This gives guidelines as to how the government and courts should treat firm (and industry) procedures such as quit fees and covenants not to compete. According to The Economist (2013) about 90 % of managerial and technical employees in the US have signed non-compete agreements, which prevents employees leaving a firm from working for a rival for a fixed period. To what extent courts honour such contracts vary. Some states in the US enforce covenants not to compete clauses in employment contracts, whereas others are more reluctant to do so. We find that ex post quit fees set by the firms improve efficiency. By contrast, covenants not to compete, interpreted as destroying matches, always harm welfare. Still firms may have an incentive to use such clauses in order to reduce worker turnover and extract worker rents ex post. Hence, it follows from our analysis that courts should be reluctant to enforce such contracts.

Related Literature There are several strands of literature that relate to our model. First, spillovers are at the core of endogenous growth models of innovation and imitation.[4] Several papers, following the seminal work by Segerstrom (1991), also analyze optimal policy.[5] However, in these

[3]Policies towards fostering innovation play an important role in many OECD countries. For instance, government-financed R&D in 2010, as a percentage of GDP, was 0.74 in the OECD and 0.92 in the US (OECD (2013)).

[4]See Eeckhout and Jovanovic (2002) and König, Lorenz, and Zilibotti (2012) for two recent examples.

[5]In particular, see Davidson and Segerstrom (1998), Mukoyama (2003), and Segerstrom (2007).

papers it is imposed by assumption that spillover effects through imitation give rise to positive externalities associated with imitation. In our model, similar effects are derived endogenously, as a result of limited commitment and search frictions in combination. Our model thus gives a microfoundation for spillover effects in labour market equilibrium.

Spillovers through worker mobility have also been studied within the industrial organization literature. Following the seminal paper by Pakes and Nitzan (1983), this literature focuses mostly on the strategic effects that arise if competitors get access to the innovation.[6] In these papers imitation is costly for the innovator because of increased price competition, not costs of replacing the worker as in our model. The inefficiency we address in this paper is related to search frictions, without such frictions our model economy is always efficient. To our knowledge, none of the imitation papers in the IO literature contains search frictions.

While our paper connects on a technical level to the literature on search with contracting under limited commitment,[7] we are not aware of any work that analyzes innovation and imitation within a labor-search environment.[8] Our mechanism, however, is related to the literature on on-the-job investments in general human capital in the presence of search frictions (see Acemoglu and Pischke (1999) and Moen and Rosén (2004)).[9] The difference between innovations and acquisition of human capital is that the latter cannot be shared costlessly. Hence if the worker obtains general human capital, a firm's output falls if a trained worker leaves and is replaced by an untrained worker. This is not the case in our innovation and imitation model. Due to the non-rivalry of information, it is sufficient that either the entrepreneur or the worker possesses information. Hence, if the employee leaves, the firm is equally well off with a new, uninformed worker, and enters the replacement market to find one. The replacement market, which is key for all our policy results, are absent in models of human capital.

[6] See also Cooper (2001), Fosfuri and Rønde (2004), Kim and Marschke (2005), and Combes and Duranton (2006).

[7] See Rudanko (2009) and Fernández-Blanco (2013).

[8] Silveira and Wright (2010) and Chiu, Meh, and Wright (2011) study the trade of knowledge in a framework with search frictions, but without looking at labor mobility. Akcigit, Celik, and Greenwood (2013) also analyse a frictional market for ideas, but their transmission mechanism is based on trade of patents. For a model of knowledge diffusion and worker mobility, where search is random and matches occur independent of equilibrium outcomes, see Lucas and Moll (2011).

[9] Marimon and Quadrini (2011) study human capital accumulation on-the-job in a setting with limited commitment, but without search frictions.

Empirical Motivation There is a substantial empirical literature that provides direct and indirect evidence on spillovers through worker flows.[10] In the following we discuss only a few of the more recent findings. First, Stoyanov and Zubanov (2012) study spillovers across firms through worker mobility by analyzing the productivity of the receiving firm measured as the value added per worker. Using Danish data they observe firm-to firm worker movements and that ”firms that hire workers from more productive firms experienced productivity gains one year after the hiring”. Greenstone, Hornbeck, and Moretti (2010) analyze productivity spillovers by comparing changes in total factor productivity of incumbent plants in a given US county stemming from the opening of new large manufacturing plants in the same county. They find that positive spillovers exist and are increasing in the worker flow between the incumbent plants' industry and the opening plants' industry.

Using similar approaches as the aforementioned studies, there is a recent strand of literature that finds evidence for labor mobility as a channel of spillovers from multinational enterprises to firms that operate only locally (see Görg and Strobl (2005), Balsvik (2011), Pesola (2011) and Poole (2013)). Further, several papers study the effect of the mobility of engineers and scientists using patent citation data and find that ideas are spread through the mobility of patent holders (see Jaffe, Trajtenberg, and Henderson (1993), Almeida and Kogut (1999), Kim and Marschke (2005), and Breschi and Lissoni (2009)). Finally, Møen (2005) finds evidence that firms use wage incentives to retain workers, who have gained knowledge of the firm's innovations, by charging a discount in the beginning of the career and paying a premium later.

The paper proceeds as follows. The two next sections (2 & 3) describe the economy and analyze the equilibrium when firms can commit to long-term wage contracts. Then, section 4 establishes efficiency of this equilibrium with full commitment. Next, section 5 studies the model when there is limited commitment on the firm side; the equilibrium is characterized and welfare properties are analyzed. After that, in the limited commitment case, welfare effects of taxes and subsidies are studied in section 6, while a detailed analysis of firm policies is undertaken in section 7. Section 8 provides a discussion of models assumptions, while section 9 concludes.

[10]There is also a large literature on productivity spillovers in general, see Bloom, Schankerman, and Van Reenen (2013) for a recent example.

2 Model Environment

There are two periods. In both periods there is a large number of potential entrepreneurs who may enter the market to start a firm. At the beginning of period 1 there is a pool of measure 1 of available workers that can be hired by entrepreneurs. The outside options of available workers are normalized to zero in both periods. As argued in the discussion section (Section 8), we may think of the set of available workers as all workers in the relevant industry. All agents are risk neutral and do not discount future values.

Production requires an entrepreneur, a worker, and knowledge. In the beginning of period 1, an entrepreneur may pay an innovation cost K to set up an innovation firm and obtain knowledge with certainty. In order to attract a worker the entrepreneur posts a vacancy at cost c with a wage contract attached to it. If a match is formed, the firm and the worker produce y per period as long as the relationship lasts. During the first period, the worker learns the innovation and becomes informed. If the entrepreneur does not attract a worker in the first period, the innovation is lost.

In the beginning of period 2, there is also entry of entrepreneurs. Instead of innovating themselves, these entrepreneurs create imitation firms, and attempt to hire a worker from an innovation firm to learn the innovation from her. This is the source of spillovers in our model. Imitation firms incur the same costs for posting vacancies (c) and have the same period output (y) as innovation firms. Finally, innovating firms that have lost their worker still possess the relevant knowledge, and can produce y with the help of a worker without knowledge. It posts a vacancy to the remaining available workers at no additional costs (for simplicity).

We assume that the economy is productive enough for there to be some entry to innovation in period 1 and some entry to imitation in period 2, while we ensure that there is no entry to innovate in period 2.[11]

Our way of modeling innovations is general enough to encompass a number of interpretations. Innovations may be technological innovations, innovations on how to use existing technology more efficiently, new management practices like lean production, new customer concepts, marketing innovations, or, in a broader setting, new product varieties (see also the discussion section). Although

[11] Parameters satisfying $2y > K + c$, $y > c$, and $y < K + c$, ensure this. Though, to ensure that variables are within the interior range of the matching function, we strengthen this to $2y(1 - \epsilon) > K + c$, $y(1 - \epsilon)^2 > c$, and $y(1 - \epsilon) < K + c$.

most of our examples involve non-patented ideas, our analysis may also be relevant for patented ideas if hiring workers from innovating firms brings spillovers that can be useful to spur new innovations.

An important assumption is that firms are small relative to the market, and only hire a limited number of workers, which in our case amounts to one. For this reason, firms in the market earn a rent, which allows them to capitalize on their initial investments. In the discussion section we argue that limited firm size may be due to decreasing returns to scale in production, or reflect that firms produce differentiated products with limited demand for each product.

We also assume that a firm has to hire a worker after it has innovated as opposed to the case where the firm innovates with an already hired worker. Our results, however, do not hinge on this timing assumption. Further, in the model economy the innovation firms cannot expand in period 2. In the discussion section we also study the effects of letting the innovation firms better exploit their own innovation by allowing them to expand by hiring more (but a limited number of) workers in period 2.

We use the search and matching technology of Diamond (1982), Mortensen (1982), and Pissarides (1985), in which a matching function maps vacancies and searching workers into a flow of new matches. Our model economy has three separate matching markets, the search market in period one, denoted by the index 1, the on-the-job search market (I), and the replacement market (R). We assume that search frictions in each market are given by the same Cobb-Douglas matching technology, $m(s_i, v_i) = As_i^{\epsilon}v_i^{1-\epsilon}$, where for each market $i \in \{1, I, R\}$, s_i and v_i are the measures of searching workers and firms with vacancies, respectively, and $\epsilon \in (0, 1)$ and $A > 0$ are parameters. Since m is the measure of matches, we require that $m(s_i, v_i) \leq min\{s_i, v_i\}$. However, we assume that the parameters of the model are such that the inequality does not bind on the relevant intervals.[12] Let $\theta_i = v_i/s_i$ denote the labor market tightness in market i. The probability of finding a worker in this market is $q(\theta_i) \equiv \frac{m(s_i,v_i)}{v_i}$, and the job finding probability is $p(\theta_i) \equiv \frac{m(s_i,v_i)}{s_i}$, implying that $p(\theta_i) = \theta_i q(\theta_i)$.

We employ the competitive search equilibrium framework of Moen (1997), where firms advertise vacancies with wage contracts attached to them, and where the wage contracts are observed by the workers before they make their search decisions. The key feature of the competitive search

[12] See the condition in footnote 11.

framework for our analysis is that it allows search externalities to be internalized. This makes it easier to identify the efficiency properties associated with the productivity spillovers. However, the competitive search framework is not crucial for our results. The important assumption is that the imitation and the replacement search markets are separate, so that the searching agents can direct their search towards the relevant market.

The following summarizes the timing protocol:

First Period:

1. Entrepreneurs enter and pay cost K in order to innovate and create an innovation firm.
2. Each innovation firm posts a wage contract at cost c to attract a worker.
3. Available workers observe the posted contracts and decide which firm to apply to.
4. Matched firms produce y units of output, while unmatched firms exit. Employed workers learn the innovation.

Second Period:

1. New entrepreneurs enter and set up an imitation firm at no costs.
2. The imitation firms post a vacancy for informed workers at cost c.
3. Innovating firms that have lost their worker post a vacancy for the remaining available workers at no costs.
4. All matched firms produce y units of output, other firms exit.

In the benchmark model we assume that firms can commit to long-term wage contracts. We then relax this assumption in section 5 and assume that firms can only commit until the end of the current period. On the workers' side we always assume lack of commitment. In particular, a worker employed at an innovation firm in the second period can break up the match before the imitation market opens, or quit to accept an offer from an imitation firm, or leave when the imitation market is closed but the replacement market is still open. Note that the participation constraint at the very beginning of period 2 (prior to entering the imitation market) can only bind if innovation firms introduce restrictions on workers who move to imitation firms, for instance in the form of a quit fee. We therefore ignore this constraint until we study restrictions on turnover in section 7.

An Excursion: The Role of Search Frictions

Before we continue, it may be enlightening to analyze the Walrasian equilibrium without search frictions and vacancy costs. The equilibrium must satisfy the zero-profit constraints of both innovation firms in period 1 and imitation firms in period 2. In addition, workers at the beginning of period 1 must be indifferent between getting a job in period 1 or waiting to get a job in the replacement market in period 2. In period 2, all workers are employed. These requirements uniquely pin down the equilibrium where: (i) a measure 1/2 of innovation firms enter in period 1 and hire half of the work force, (ii) a measure 1/2 of imitation firms enter the market in period 2 and hire all employed workers, and (iii) the innovation firms hire all the remaining available workers in the replacement market.[13]

On average, the worker works in 3/2 periods and produces y per period, and the investment cost per worker is $K/2$. The total wage income over the two periods is then $y3/2 - K/2$. If imitation was impossible, all workers would be hired in period 1, and the total wage income would be $2y - K$. Hence, the gain from imitation is $(K - y)/2 > 0$. It is easy to verify that the Walrasian equilibrium allocation is efficient. This allocation emerges independently of the assumptions made on commitment of innovation firms, as competition between imitation firms always increases the wage paid by imitation firms up to y.

3 Model with Full Commitment

This section analyzes the model where firms can fully commit to long-term wage contracts.[14]

For an employed worker in period 1, the value of a contract at the beginning of period 1 is given by

$$W_1 = w_1 + W_2, \tag{1}$$

where w_1 is the period-1 wage offered by an innovating firm. The value of the contract at the beginning of period 2 is given by

$$W_2 = p(\theta_I)w_I + (1 - p(\theta_I))w_2, \tag{2}$$

[13] The wage structure supporting this equilibrium is $w_1 = \frac{1}{2}y - \frac{1}{2}K$, $w_I = y$, and $w_R = \frac{3}{2}y - \frac{1}{2}K$, where w_1 denotes the period-1 wage, w_I the wage paid by imitation firms, and w_R the wage paid by innovation firms to their new hires in period 2. Note, w_1 is negative since $K > y$.

[14] To be precise, we consider one-sided full commitment, since we assume throughout this paper that workers do not commit.

where w_2 is the period-2 wage offered by an innovating firm, w_I is the wage offered by an imitating firm in period 2, and $p(\theta_I)$ is the probability of finding a job at an imitating firm.[15] That is, the value of a worker in an innovating firm at the beginning of period 2 is the promised wage w_2 plus the the expected surplus of searching for a job at an imitating firm.

The values of an available worker at the beginning of period 1 and period 2 are

$$U_1 = p(\theta_1)W_1 + (1 - p(\theta_1))U_2 \tag{3}$$

and

$$U_2 = \ p(\theta_R)w_R, \tag{4}$$

respectively, where w_R is the wage offered in the replacement market, and $p(\theta_1)$ and $p(\theta_R)$ are the job finding probabilities in the period-1 hiring market and the replacement market, respectively. Recall that a worker that remains unmatched receives a period income normalized to zero.

The profit of an innovating firm in period 1 that has already hired a worker is given by

$$J_1 = y - w_1 + p(\theta_I)V_R + (1 - p(\theta_I))(y - w_2), \tag{5}$$

where V_R is the value of a vacancy posted in the replacement market, given by

$$V_R = q(\theta_R)(y - w_R), \tag{6}$$

where θ_R is the labor market tightness in the replacement market. Note that there is no free entry in the market for replacement workers in the second period, since only innovating firms that have already entered in the first period can post vacancies. Therefore, the market tightness θ_R is completely determined by the market tightness of the other markets. Since the mass of workers in the economy is one, we have

$$\theta_R = \frac{p(\theta_1)p(\theta_I)}{1 - p(\theta_1)}, \tag{7}$$

where the numerator is derived from the fact that the measure workers of innovating firms that have lost their employee at the beginning of period 2 equals the number of workers who have found a job at an imitation firm. The ex-ante value of innovating and opening a vacancy in an innovation firm is

$$V_1 = q(\theta_1)J_1 - c - K, \tag{8}$$

[15]If $w_I \leq w_2$, workers will not search, and $p(\theta_I) = 0$.

where $q(\theta_1)$ is the probability that the vacancy is filled. The value of a vacancy in an imitation firm is

$$V_I = q(\theta_I)(y - w_I) - c, \tag{9}$$

where $q(\theta_I)$ is the job-filling probability.

In our competitive search environment firms post a wage contract that is observable to all workers. When deciding on wages, a firm has rational expectations on how the value of the wage contract offered influences its chance of hiring a worker, or, equivalently, the labour market tightness (or the inverse of the worker queue length) they will face. A higher expected value of the contract, implies a higher probability of hiring a worker. Search is competitive as all firms have to offer an expected value of search that is no lower than the expected value workers could get elsewhere in the market. That is, when hiring, innovating firms take the equilibrium values U_1^* and U_2^* as given, while imitation firms take the equilibrium value W_2^* as given. It then follows that equations (2), (3), and (4) are constraints for the firms that form relationships between advertised values of contracts and labour market tightnesses.[16]

In addition to the standard assumptions regarding advertised wages and the probability of hiring workers, innovation firms also have to form expectations about the relationship between the period-2 wage w_2 they offer to the worker and the probability $p(\theta_I)$ that the worker quits. We follow here the literature on competitive search on-the-job (see Moen and Rosén (2004), Shi (2009), and Menzio and Shi (2010)). Suppose a small subset of innovating firms offer a wage $\bar{w}_2$, which may be different from the equilibrium wage. Then a submarket opens up, and imitating firms flow into this submarket up to the point where they receive zero profits. They offer wages $w_I(\bar{w})$ so as to maximize profit, taking the expected market value of search of the workers in this submarket as given. It follows that the resulting values of θ_I and w_I, denoted by and $\widehat{\theta}_I(\bar{w}_2)$ and $\widehat{w}_I(\bar{w}_2)$, are given by[17]

$$\{\widehat{\theta}_I(\bar{w}_2), \widehat{w}_I(\bar{w}_2)\} = \underset{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I = 0}{\arg\max} \quad p(\tilde{\theta}_I)\tilde{w}_I + (1 - p(\tilde{\theta}_I))\bar{w}_2 \tag{10}$$

The assumption is that, when deciding on w_2, workers and firms alike expect that workers will quit and start in an imitation firm and receive a wage $\widehat{w}_I(w_2)$ with probability $\widehat{p}_I(w_2) \equiv p(\widehat{\theta}_I(w_2))$. It

[16]In addition there are the following participation constraints for the first and the second period, respectively: $U_2 \leq W_1$ and $U_2 \leq W_2$. One can show that they do not bind in equilibrium.

[17]This is the dual problem of profit maximization subject to the zero-profit condition.

follows that we can write

$$V_1 = q(\theta_1)[y + \widehat{p}_I(w_2)V_R + (1 - \widehat{p}_I(w_2))(y - w_2)] - c - K, \tag{11}$$

$$W_1 = w_1 + \widehat{p}_I(w_2)\widehat{w}_I(w_2) + (1 - \widehat{p}_I(w_2))w_2. \tag{12}$$

3.1 Equilibrium

Definition 1 *An equilibrium is a vector of market tightnesses* $\{\theta_1^*, \theta_I^*, \theta_R^*\}$, *values for workers* $\{W_1^*, W_2^*, U_1^*, U_2^*\}$, *and values for firms* $\{V_1^*, V_I^*, V_R^*\}$ *(where all values are according to the definitions above), a contract* $\{w_1^*, w_2^*\}$, *and wages* $\{w_I^*, w_R^*\}$ *satisfying the following conditions:*

1. *Optimal Contract and Profit Maximization:*

   (a) The contract $\{w_1^*, w_2^*\}$, *maximizes* V_1 *given by (11) subject to (3) and (12);*

   (b) The wage w_I^* *maximizes* V_I *given by (9) subject to (2);*

   (c) The wage w_R^* *maximizes* V_R *given by (6) subject to (4).*

2. *Zero-Profit Conditions:* $V_1^* = V_I^* = 0$.

3. *The labor market tightness in the replacement market,* θ_R^*, *is given by (7).*

3.2 Characterization of Equilibrium

We start with the period-2 decisions to solve for equilibrium. First, consider the imitating firm's problem of maximizing V_I given by (9) subject to (2). The optimal wage conditional on w_2 is given by[18]

$$\widehat{w}_I(w_2) = \epsilon y + (1 - \epsilon)w_2. \tag{13}$$

This is the standard result in competitive search models: the surplus (here $y - w_2$) is shared between the worker and the firm according to the elasticity of the job finding probability, i.e. ϵ. Then, by using (13) to substitute out $\widehat{w}_I(w_2)$ in (9), the zero-profit condition for the imitating firms implicitly determines $\widehat{\theta}_I(w_2)$:

$$q(\widehat{\theta}_I(w_2)) = \frac{c}{(1 - \epsilon)(y - w_2)}. \tag{14}$$

Given the solution for $\widehat{\theta}_I(w_2)$, we then have $\widehat{p}_I(w_2)$.

[18]A derivation of the first order condition is given in appendix 10.1.

Next, consider the replacement market in period 2. The innovation firm sets w_R so as to maximize V_R given by (6) subject to (4), with first order condition

$$w_R = \epsilon y, \tag{15}$$

independently of θ_R. Given θ_R, which is determined by the tightness in the other markets, this pins down V_R and U_2:

$$\begin{aligned} V_R &= q(\theta_R)(1-\epsilon)y \\ U_2 &= p(\theta_R)\epsilon y. \end{aligned} \tag{16}$$

We now turn to the innovating firm's problem in period 1. It is instructive to divide this maximization problem into two steps:

1. *Optimal retention*: For a given W_1, find the contract $\{w_1, w_2\}$ that maximizes J_1 given the functions $\widehat{p}_I(w_2)$ and $\widehat{w}_I(w_2)$.

2. *Optimal recruiting*: Find the value of W_1 that maximizes V_1 subject to the constraint (3).

Before we proceed, define $M_i \equiv W_i + J_i$ to be the joint income of a matched worker and firm in period i. M_2 can be written as

$$M_2 = y + \widehat{p}_I(w_2)[V_R + \widehat{w}_I(w_2) - y]. \tag{17}$$

The joint income of a matched worker-firm pair in period 2 is then given by

$$M_1 = y + M_2. \tag{18}$$

To solve for step 1, we first rewrite (5) as

$$\begin{aligned} J_1 = M_1 - W_1 &= 2y + \widehat{p}_I(w_2)[V_R + \widehat{w}_I(w_2) - y] - W_1 \\ &= 2y + \widehat{p}_I(w_2)[V_R + w_2 - y] + \widehat{p}_I(w_2)(\widehat{w}_I(w_2) - w_2) - W_1. \end{aligned} \tag{19}$$

The first order condition with respect to w_2 can then be written

$$\begin{aligned} \frac{dJ_1}{dw_2} &= \frac{d\widehat{p}_I(w_2)}{dw_2}[V_R + w_2 - y] + \widehat{p}_I(w_2) + \frac{d}{dw_2}[\widehat{p}_I(w_2)(\widehat{w}_I(w_2) - w_2)] \\ &= 0. \end{aligned} \tag{20}$$

Remember that the dual problem (10) implies that $\{\widehat{\theta}_I(\bar{w}_2), \widehat{w}_I(\bar{w}_2)\}$ solve $\max_{\{\tilde{\theta}_I, \tilde{w}_I\}} p(\tilde{\theta}_I)(\tilde{w}_I - w_2)$ subject to $V_I = 0$. Hence, it follows from the envelope theorem that $\frac{d}{dw_2}[\widehat{p}_I(w_2)(\widehat{p}_I(w_2) - w_2)] = -\widehat{p}_I(w_2)$. Thus, the first order condition with respect to w_2 reduces to[19]

$$w_2 = y - V_R. \tag{21}$$

This expression says that the worker in period 2 gets all the value created in period 2 net of the expected profits of the firm from hiring in the replacement market. At this wage, the worker is the sole residual claimant, and thus takes into account the full opportunity costs of leaving to the imitation firm. This implies that the worker's on-the-job search decision exerts no negative externality on the firm, and hence joint income is maximized. Although the firm receives zero net profit in the second period, it can extract surplus from the worker in period 1 through w_1.

Turning to the optimal recruiting problem in step two, the firm now takes M_1 as given and maximizes $V_1 = q(\theta_1)(M_1 - W_1) - c - K$ subject to (3). The first order condition

$$W_1 = \epsilon M_1 + (1 - \epsilon) U_2, \tag{22}$$

gives that the value of the contract offered by the firm is a share of the match surplus $(M_1 - U_2)$.

By substituting in all the equilibrium values into (8), the zero-profit condition for innovating firms becomes:

$$V_1 = q(\theta_1) y (1 - \epsilon)[2 + p(\theta_I) q(\theta_R) \epsilon (1 - \epsilon) - \epsilon p(\theta_R)] - c - K = 0. \tag{23}$$

Similarly, by substituting equilibrium values into into (9), we obtain for imitating firms:

$$V_I = q(\theta_I) q(\theta_R) y (1 - \epsilon)^2 - c = 0. \tag{24}$$

Given the definition of $\theta_R \equiv \frac{p(\theta_1) p_I(\theta_I)}{1 - p(\theta_1)}$, the two equations (23) and (24) determine the equilibrium allocation $\{\theta_1^*, \theta_I^*\}$. In appendix 10.3 we show the following result:

Proposition 1 *An equilibrium exists and is unique.*

Uniqueness is not trivial due replacement-market effects. To see the main step of the proof, note that by using the results of the optimal recruiting problem we can rewrite the maximized value

[19] See appendix 10.2 for more details.

of V_1 as[20]

$$V_1(\theta_1) = q(\theta_1)(1-\epsilon)[2y + \max_{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I=0} \{p(\tilde{\theta}_I)[V_R + \tilde{w}_I - y]\} - U_2] - c - K,$$

where the maximization problem gives the workers' gain from search in equilibrium, with $y - V_R$ inserted for w_2. We have to show that this is strictly decreasing in θ_1. An increased value of θ_1 influences the value of innovation through several channels. First, it makes it more costly to find a worker in period 1, which tends to reduce V_1. Second, a higher value of θ_1 makes the replacement market tighter, and thereby reduces V_R and U_2 and hence also V_1. The third effect is an indirect effect. A lower value of V_R increases w_2 and therefore reduces worker turnover $p(\theta_I)$ and increases w_I. However, due to efficient contracting (the firm is indifferent as to whether the worker quits) and the envelope theorem (as θ_I and w_I maximize the searching workers' expected income given the zero-profit constraint of imitation vacancies) these indirect effects have no impact on V_1. It follows that the value of innovating unambiguously falls with θ_1, and this ensures that the equilibrium is unique.

4 Efficiency

In this section we determine the constrained efficient allocation and compare it to the equilibrium allocation of the full commitment case.

As it is common in the literature, we measure welfare as total output net of innovation and vacancy costs. By constrained efficiency we mean that the social planner faces the same matching frictions as the agents in the market. Since the mass of available workers is normalized to unity,[21] aggregate output in period 1 equals $p(\theta_1)y - \theta_1(c+K)$. If a worker at an innovation firm moves to a imitation firm in period 2, her contribution to output is unchanged. However, the now vacant innovation firm will produce additional output only if it is able to hire a new worker. Aggregate net output therefore is

$$F(\theta_1, \theta_I) = p(\theta_1)[2y + p(\theta_I)q(\theta_R)y - c\theta_I] - (c+K)\theta_1, \tag{25}$$

[20]Here, we express the firm's maximization of M_1 as a choice directly over θ_I and w_I instead of a choice over $\widehat{p}_I(w_2)$ and $\widehat{w}_I(w_2)$ by setting w_2.

[21]This implies both that $p(\theta_1)$ is equal to number of workers that find a job and that θ_1 is equal to the number of vacancies in the first period.

where, as before, $\theta_R = \frac{p(\theta_1)p_I(\theta_I)}{1-p(\theta_1)}$. The planner chooses θ_1 and θ_I so as to maximize welfare. The first order condition for θ_1, after some manipulation, can be written as[22]

$$\frac{\partial F}{\partial \theta_1} = q(\theta_1)y(1-\epsilon)[2 + p(\theta_I)q(\theta_R)\epsilon(1-\epsilon) - \epsilon p(\theta_R)] - c - K = 0. \tag{26}$$

The first order condition with respect to θ_I is

$$\frac{\partial F}{\partial \theta_I} = p(\theta_1)[q(\theta_I)q(\theta_R)y(1-\epsilon)^2 - c] = 0. \tag{27}$$

Comparing these two first order conditions to the zero-profit conditions in equilibrium for innovators (23) and for imitators (24), it is immediate that the (necessary) equilibrium conditions are identical to the necessary conditions for the interior efficient allocation. Thus we have established the following result:

Proposition 2 *The full-commitment equilibrium allocation is constrained efficient.*

Efficiency in the commitment case can be explained by contracting under full commitment and competitive search. The ument can be divided into several steps.

First, the on-the-job search market in period 2 maximizes the income of the searching worker given the constraint that the imitation firms must make zero profits. Hence, the worker receives the entire social gain from her knowledge about the innovation. Second, when the worker searches so as to maximize her own income in period 2, there are no externalities from her search behavior on the employer. The period-2 wage in the innovation firm is exactly equal to the opportunity cost of letting the worker move to an imitation firm, i.e. output less the value of a vacancy in the replacement market. Thus, when maximizing her own income, the worker in effect also maximizes joint income. Third, the firm commits to a total compensation value at the beginning of period 1. The worker therefore only cares about the total compensation and will accept a low wage in period 1. Thus the firm can extract the value of the innovation net of the total wage costs. Finally, innovating firms compete for the workers ex ante, and enter up to the point where the gain from entering is equal to the cost. Since search is competitive, this process does not create distortions, and efficiency prevails.

[22]See appendix 10.4 for details of the algebra.

To sum up, the optimal decision for the firm is to give the full income to the worker in period 2, and extract income only in period 1 through w_1. Joint income maximization implies that also the worker's surplus is maximized, i.e. the worker will search optimally, which is efficient from the social planner's point of view.

5 Model with Limited Commitment

The fact that the equilibrium analyzed above is efficient rests on the firms' commitment to future wages. Empirically however, such wage commitments seem to be a strong assumption, since firms often fire workers at low costs (see Boeri, Garibaldi, and Moen (2013)). Also from a theoretical point of view firms may be reluctant to commit to wages. In a broader setting there may be contingencies in which a firm would like to reduce the wage or even fire the worker, and these contingencies may be hard to prove in a court of law. For instance, the firm may wish to have the option to cut wages if workers perform badly. Hence, the firm may want to have some flexibility regarding wage setting ex post, despite the fact that this may lead to inefficiencies along other dimensions like optimal on-the-job search.

This section analyzes the model where firms can only commit to the wage within the current period. We call this the case of limited commitment. The model is identical to the full commitment case except for the determination of w_2. We consider two wage setting procedures which are both commonly used in the labor-search literature:

1. Wage posting: Firms set the wage before the workers make their on-the-job search decisions (denoted by subscript A).

2. Wage bargaining: Workers and firms bargain over wages after the on-the-job search market closes, but before the replacement market closes (denoted by subscript B).

We will discuss both wage setting procedures in turn. In the wage posting case the period-2 profit of an innovating firm can be written as

$$J_2 = y - w_{2A} + \widehat{p}_I(w_2)[V_R + w_{2A} - y],$$

where $\widehat{p}_I(w_2)$ is defined by (10) as before. The firm optimally sets w_{2A} by anticipating how this wage will affect the probability that the worker leaves to an imitating firm:

$$\begin{aligned} &\max_{\tilde{w}_{2A}} J_2 \\ \text{s. to} \quad &\tilde{w}_{2A} \geq U_2. \end{aligned} \tag{28}$$

In appendix 10.5 we show that

$$w_{2A} = \max\{y - \frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon} V_R, U_2\}. \tag{29}$$

A necessary condition for an interior solution is that $p(\theta_I) > \epsilon$, otherwise the wage floor U_2 will bind.[23] Given an interior solution, and the fact that $\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon} > 1$, it follows that $U_2 \leq w_{2A} < y - V_R$.[24]

For the bargaining case, we employ the Nash sharing rule, with V_R being the firm's outside option and U_2 being the worker's outside option. As is standard in matching models, we assume that the outside options are also the threat points. The Nash problem then has the following standard solution

$$w_{2B} = \beta(y - V_R) + (1-\beta)U_2, \tag{30}$$

where $\beta \in (0,1)$ is the worker's bargaining power. Note that $U_2 < w_{2B} < y - V_R$. Further, note that in this case, the imitating firm anticipates the wage the innovating firm will bargain for if it remains with its worker.

In both cases, the wage depends only on θ_R in equilibrium.[25] Hence, in general we can write

$$w_2 = w_{2j}(\theta_R), \tag{31}$$

where $j = A$ if wages are determined by wage posting and $j = B$ if wages are determined by bargaining. For both wage setting procedures it holds that $\frac{\partial w_{2j}(\theta_R)}{\partial \theta_R} > 0$. Further, as has been noted above, in both cases, the firm pays a second-period wage that is below the net value of searching for a replacement worker, $y - V_R$. In the wage posting case, this is because the firm now trades off the retention and the rent extraction within the period, while in the case of bargaining, the outcome directly follows from the Nash sharing rule.

[23] The second order conditions are derived in 10.5.

[24] $U_2 < y - V_R$ follows from the definitions of the terms together with the fact that $y > \epsilon y \equiv w_R$.

[25] In the wage bargaining case, this is obvious. With wage posting, note that we can write $w_{2A} = f(\theta_I, \theta_R)$ from the wage equation and $\theta_I = g(w_{2A})$ from the zero profit condition for imitating firms, where f and g are both continuous, $f_{\theta_I}, f_{\theta_R} \geq 0$, and $g'(w_2) < 0$. For a given θ_R, w_{2A} solves $f(g(w_2), \theta_R) - w_{2A} = 0$. It is straightforward to show that this equation has a solution, which we denote $w_{2A}(\theta_R)$, and that the solution is well behaved and increasing in θ_R. Note that $w_{2A}(\theta_R)$ is derived under the assumption that the zero-profit constraint on imitation firms is satisfied.

Finally, we incorporate the possibility that innovation firms can match wage offers. With limited commitment, we cannot ex ante rule out that $w_I < y - V_R$, in which case it would be in the interest of the innovating firm to match the wage offer. If we allowed for offers and counter-offers, the wage w_I would be driven up to $y - V_R$, since this is the value the innovating firm would get from the replacement market. Instead of modeling the offer game explicitly, we assume that the imitating firm's wage has to satisfy the following wage floor constraint: $w_I \geq y - V_R$. Technically, imitation firms choose the wage w_I so as to maximize V_I given by (9) subject to (2) and $w_I \geq y - V_R$.

In period 1, innovation firms choose w_1 so as to maximize V_1 given by (11) subject to (3) and (12), with w_2 given by (31), and taking into account the new constraint on imitation firms.

In the appendix 10.6 we show the following proposition:

Proposition 3 *The limited commitment equilibrium exists.*

Since we cannot rule out that the value of operating an innovation firm in period 2 may be increasing in θ_1, uniqueness may not be guaranteed. To see this, consider period-2 profits, given by $J_2 = y - w_2 + p(\theta_I)[V_R + w_2 - y]$. Due to the fact that $w_2 < y - V_R$ (and hence $y - w_2 > V_R$), J_2 is decreasing in $p(\theta_I)$. A higher θ_1 means a lower θ_I, which will tend to increase J_2. On the other hand, as ued in the proof of proposition 1, V_R decreases if θ_1 increases. Hence, we cannot rule out that J_2 is increasing in θ_1. In principle this may give rise to multiple equilibria. If so, the equilibria can be welfare ranked; the equilibrium with the highest θ_1 gives the highest U_1 and, hence, also highest welfare. If the model has multiple equilibria, we assume that the agents coordinate on the equilibrium with the highest θ_1.

5.1 Equilibrium and Welfare

As shown above, limited commitment will lead to a lower w_2 in equilibrium compared with the full commitment case. The lower wage, in turn, will lead to a higher probability of losing the worker to an imitating firm. The total effect is that the joint income of a matched worker-firm pair in period 1 is lower, and, hence, also θ_1 is lower.

Proposition 4 *The limited-commitment equilibrium has higher θ_I and lower θ_1 than the full-commitment equilibrium.*

Proof. See appendix 10.6 ■

The limited-commitment allocation is clearly not efficient since it differs from the unique efficient allocation under full commitment:[26]

Corollary 1 *The limited-commitment allocation is not constrained efficient.*

The intuition for the inefficiency result is as follows. First, the wage w_2 will be lower than in the full commitment case. As a result, too many imitation vacancies are posted, paying too low wages. Put differently, the imitation market maximizes the income of the searching workers (given the zero-profit constraint of imitation vacancies). Since w_2 is too low, quits impose a negative externality on the employers, and the joint income of an innovation firm and its employee is lower than what it would have been if innovating firms were setting a higher wage with a corresponding lower $p(\theta_I)$. In period 1, the innovation firm may still extract the period 2 surplus from the worker, but joint income is smaller than in the full commitment case. As a result, fewer innovation firms enter the market, and welfare is lower.

To gain more insight into the inefficiency result, we continue by analyzing the welfare function evaluated at the limited-commitment allocation. Recall that the aggregate output in the economy, absent any policy, is given by $F(\theta_1, \theta_I)$ defined in (25). Let θ_1^{**} and θ_I^{**} denote the limited-commitment equilibrium values of θ_1 and θ_I, respectively. Then the following holds:

Lemma 1 *The following conditions are satisfied at the limited-commitment allocation:*

$$\frac{\partial F(\theta_1^{**}, \theta_I^{**})}{\partial \theta_1} = 0$$
$$\frac{\partial F(\theta_1^{**}, \theta_I^{**})}{\partial \theta_I} < 0.$$

Proof. See appendix 10.7. ■

The excessively high equilibrium value of θ_I reduces the magnitude of θ_1 required for optimal first-period entry compared to the full commitment level. However, given θ_I^{**} the level of θ_1^{**} is welfare maximizing. In contrast, a marginal reduction in θ_I from its limited-commitment value is strictly welfare improving. This result is very helpful for policy analysis. It implies that for any

[26]Note that the excluded boundary value $\beta = 1$ for the bargaining case would imply the efficient allocation. Giving the worker the full bargaining power in period 2 is equivalent to the outcome in the full commitment model, where the innovating firm chooses to give all the second period surplus to its worker.

policy that does not alter the welfare function F itself, we know that the policy is welfare improving if and only if it reduces θ_I. Monetary transfers between the agents will not affect the structure of F. However, policies that involve real costs (like an increase of the matching friction in the on-the-job search market) will.

6 Taxes and subsidies

In the equilibrium with limited commitment there is too little innovation and too much imitation compared with the full commitment case. This inefficiency gives scope for welfare improving policies. As outlined in the introduction, policies to foster innovation and restrictions on worker mobility to curb imitation are widespread and substantial. Since our model makes the transmission mechanism of productivity spillovers explicit, our analysis not only determines the resulting welfare effects, but also illuminates the way these policies function.

In this section we analyze subsidies to innovation and taxes on imitation, while section 7 analyzes policies that extend the contracting possibilities of the firms.

Define $s > 0$ as a subsidy to vacancy creation in period 1 and $\tau > 0$ as a tax on vacancy creation in period 2.[27] We assume that any net receipts or losses are redistributed in a lump-sum fashion to all workers equally.

It is easy to show that there exist a tax on imitating firms together with a subsidy to innovating firms that lead to the efficient allocation. A subsidy will induce entry in period 1 so that θ_1 increases. At the optimal θ_1, we know that θ_I is too high in the limited commitment case. Thus a tax is needed in addition to the subsidy to obtain efficiency.

Next we analyze the two instruments in isolation. Beginning with the subsidy, we have the following result:

Proposition 5 *A subsidy to the innovating firms will increase the number of innovating firms, will decrease the tightness in the imitation market θ_I, and will increase welfare.*

Proof. See appendix 10.8. ■

Intuitively, a subsidy will directly increase the number of innovating firms, and thereby θ_1. Higher θ_1 will increase the cost of replacement and thereby increase the wage w_2. A higher wage

[27] For most of the analysis we focus on lump-sum taxes and subsidies on ex-ante profits. However, it is clear that there exist linear taxes and subsidies on ex-post profits that can achieve the same result.

then reduces the imitation rate. It is this induced negative effect of θ_1 on θ_I that raises the level of welfare. Formally,

$$\frac{dF}{d\theta_1} = \frac{\partial F}{\partial \theta_1} + \frac{\partial F}{\partial \theta_I}\frac{d\theta_I}{d\theta_1} = \frac{\partial F}{\partial \theta_I}\frac{d\theta_I}{d\theta_1} > 0.$$

Thus, only the effect through θ_I is relevant for the result. This mechanism highlights the role of the general equilibrium effect coming from the replacement market: if the expected value of replacing a worker, V_R, would be unaffected by labor market conditions, a subsidy would have no effect on welfare.

Turning now to the tax on imitation, we have the following:

Proposition 6 *A tax on imitating firms profits will reduce θ_I and will increase welfare.*

Proof. See appendix 10.9. ■

A tax on imitation firms reduces the leaving rate of workers from innovation firms to imitation firms. This effect directly increases welfare. However, whether a tax on imitation spurs or hurts innovation, i.e. whether θ_1 increases or decreases, is less clear. The reason is that the tax redistributes surplus and thus influences period 2 profits of innovating firms. On the one hand, a lower imitation probability may reduce the joint income of the innovating firm and its employee. On the other hand, less imitation also reduces U_2 and thus also the share of the joint income allocated to the worker. Hence, whether a lower θ_I increases period 2 profits of innovating firms is unclear and may depend on parameters. Regarding the effect on welfare, however, lemma 1 implies that this ambiguity is without consequence since $\partial F/\partial \theta_1 = 0$.

However, if the imitating firm's wage is at the floor, i.e. $w_I = \check{w}_I \equiv y - V_R$, we get a clear effect of the tax on innovation. In this case the ex-post profits of the innovating firm in equilibrium become:

$$J_1 = (1-\epsilon)[2y + p(\theta_I)(V_R - (y - w_I)) - U_2] = (1-\epsilon)[2y - U_2],$$

i.e. the joint income M_1 is a constant equal to $2y$. Thus, profits depend on θ_I only through U_2, which is decreasing in θ_I. Hence, in this case a tax on imitation will increase entry to innovation, θ_1.

7 Firm policies

In this section we extend the contractual toolbox of the firm. We study the incentives of firms to employ more sophisticated contracts, as well as their welfare implications. In particular we study the effects of quit fees, covenants not to compete, and options of renting out workers. Our analysis gives useful guidelines regarding the attitude the government should take towards these firm policies.

Quit fees Restrictions on worker mobility may take different forms. The milder form is that the worker has to compensate the firm if the worker leaves. We refer to this as a quit fee. A more drastic measure is that firms (or an employer association in an industry) introduce covenants not to compete, restricting movements of workers between firms through "brute force". In this subsection we consider a quit fee α paid by the worker to the innovating firm if she leaves.

If the firm can commit to α at the hiring stage, efficiency will, not surprisingly, be restored. One can easily show that the innovation firm can influence the search behavior of the worker in period 2 through its choice of α. Maximizing joint income with respect to α will then be a substitute for maximizing with respect to w_2.[28]

Assume now instead that the firm cannot commit to α in period 1, but sets α at the beginning of period 2. More precisely, the firm posts contract $\{w_2, \alpha\}$, which the worker accepts or rejects. This reduces the value of the worker of being employed in the innovation firm when searching for a job in an innovation firm. Recall that the worker may leave the firm at will at the beginning of period 2, before searching for a job in an imitation firm. The associated participation constraint, until now always satisfied, may thus bind. We refer to this constraint, somewhat imprecisely, as the interim participation constraint.

The outside option of the worker at this stage reads

$$\overline{W} \equiv \max_{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I=0} [p(\tilde{\theta}_I)\tilde{w}_I + (1 - p(\tilde{\theta}_I))U_2]. \tag{32}$$

Even if the firm has all the bargaining power, the interim participation constraint implies that

[28] It does so by influencing w_I through α. The ex-post payment to the worker when she pays a quit fee is $w_I = \epsilon y + (1-\epsilon)(w_2 + \alpha)$. Then, together with the fact that there is a lower bound on w_2, it is clear that by choosing α the firm can set w_I equal to the efficient level w_I^* for any given level w_2.

the firm has to offer a contract that satisfies $W_2 \geq \bar{W}$. Notice that for any $\alpha > 0$ the interim participation constraint implies that $w_2 > U_2$.

The expected period-2 income of the worker, if entering the period as employed by an innovation firm, is $W_2 = w_2 + p(\theta_I)(w_I - \alpha - w_2)$. Analogous to (10), it follows that the values $\widehat{\theta}_I$ and $\widehat{w}_I$, as implicit functions of α and w_2, maximize W_2 given the zero profit constraint of imitation firms:

$$\{\widehat{\theta}_I, \widehat{w}_I\} = \underset{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I = 0}{\arg\max} \ [w_2 + p(\tilde{\theta}_I)(\tilde{w}_I \ - \alpha - w_2)]. \tag{33}$$

Clearly, $\widehat{w}_I$ and $p(\widehat{\theta}_I)$ only depend on the sum $\alpha + w_2$. The firm maximizes ex post profits, J_2, given by

$$J_2 = p(\hat{\theta}_I(w_2 + \alpha))(\alpha + V_R) + (1 - p(\hat{\theta}_I(w_2 + \alpha)))(y - w_2), \tag{34}$$

with respect to w_2 and α, subject to (32).

The first thing to note is that the constraint (32) always binds. If not, the innovating firm could lower w_2, and at the same time increase α by the same amount. This would not influence $p(\widehat{\theta}_I)$. However, the firm's ex post profit would increase. Substituting (32) (which binding so that $W_2 = \overline{W}$) into the expression for J_2 from above gives

$$\begin{aligned} J_2 &= p(\hat{\theta}_I)(\hat{w}_I + V_R) + (1 - p(\hat{\theta}_I))y - \bar{W} \\ &= M_2 - \bar{W}, \end{aligned} \tag{35}$$

where M_2 is the joint income as defined in equation (17). This is parallel to the first step of the firm's maximization problem in the full commitment case, where the firm maximizes $M_1 - W_1$ with respect to w_2 for W_1 given. To be more precise, the problem of maximizing $M_2 - \bar{W}$ given by (35) with respect to w_2 is equivalent to the problem of maximizing M_1 given by (19) with respect to w_2 up to a constant, hence the two problems have the same solution. In both cases, the firm is the residual claimant, and thus has an incentive to maximize joint income. The firm induces optimal on-the-job search by setting $w_2 + \alpha = y - V_R$.

To complete the analysis, insert the first order condition for w_I (analogous to (13), given by

$w_I = \varepsilon y + (1-\varepsilon)(w_2+\alpha))$, and $w_2 + \alpha = y - V_R$, into the expression for W_2 to obtain

$$\begin{aligned} W_2 &= w_2 + p(\widehat{\theta}_I)(\epsilon y + (1-\varepsilon)(y - V_R) - (y - V_R)) \\ &= w_2 + \epsilon V_R. \end{aligned}$$

The value of w_2 then solves $W_2 = \overline{W}$. We have the following proposition:

Proposition 7 *If the firm can post a contract in the second period specifying a quit fee α and a wage w_2, the efficient allocation is attained. The wage w_2 is lower than in the full commitment case.*

Efficiency is obtained because with the quit fee the firm has two instruments. This enables the firm to both extract all the rent from the worker, and in addition govern her search behavior. As a result, the trade-off between rent extraction and efficiency is defused, the firm becomes the residual claimant and implements efficiency.

Compared with the full-commitment case, the wage profile is more front-loaded with limited commitment and quit fees. Both workers and firms realize ex ante that the firms will extract rents ex post, and as a result there is fiercer competition leading to higher wages paid in period 1.

It follows from our analysis that allowing the firm to charge a quit fee restores efficiency, even if it is agreed upon ex post, and hence that such arrangements should be approved by a court of law. However, we have one caveat here, as the argument rests on the presumption that the workers ex ante anticipate that they will have to pay a quit fee if they find a new job ex post. If workers do not anticipate this, their wages will be lower than expected, and too many innovation firms will enter in period 1.

Covenants Not to Compete We now turn to covenants not to compete as a firm policy. In this subsection we assume that firms cannot enforce a quit fee if the worker leaves, but it can restrict the movement of the worker. The restriction, through clauses in the work contract or through industry standards, makes it more difficult for the worker to search on-the-job or harder to change jobs once a job is found (for instance because of possible law suits).

In our model, these types of restrictions on mobility can be interpreted as less efficient hiring, that is, a reduction in the number of matches for a given market tightness. More concretely, now the probability of finding a worker for imitating firms is given by $(1-\rho)q(\theta_I)$ (and the job finding probability in the imitation market by $(1-\rho)p(\theta_I)$), where $\rho \in [0,1]$ is a measure of the strictness of the covenants not to compete.

To understand the welfare effects of covenants not to compete, let us first derive the planner's choice of ρ. More specifically, we write the matching function as $(1-\rho)m(s,v)$, and let the planner decide on ρ. For a given ρ, the equilibrium is defined as above. Note that, for a given ρ, the matching function is well defined, and the welfare function (25) becomes

$$F(\theta_1, \theta_I, \rho) = p(\theta_1)[2y + (1-\rho)p(\theta_I)q(\theta_R)y - c\theta_I] - (K+c)\theta_1.$$

The first thing to note is that ceteris paribus, for given values of θ_I and θ_1, an increase in ρ decreases welfare. However, an increase in ρ will change the equilibrium values of θ_1 and θ_I. For a given ρ, the matching function is well defined, and lemma 1 holds. Hence, we know that an equilibrium response in θ_1 has no effect on welfare. However, equilibrium effects on θ_I do have welfare consequences. If θ_I increases, this will reduce welfare even further. However, if θ_I decreases, this will tend to increase welfare, and the net effect is not obvious. We want to show that welfare decreases also in this case. It turns out that this is rather an intricate problem, as there are many effects, and in addition, we have to take into account the lower bound on w_I ($w_I \geq y - V_R$). It will prove to be convenient to substitute w_I back into the welfare function. Recall that in equilibrium $(1-\rho)q(\theta_I)(y-w_I) = c$. Using this and the relationship $\theta_I q(\theta_I) = p(\theta_I)$, it follows that $c\theta_I = (1-\rho)q(\theta_I)(y-w_I)$. If we substitute this into the expression for F, we find that the equilibrium welfare as a function of ρ writes

$$\hat{F}(\theta_1(\rho), \theta_I(\rho), w_I(\rho), \rho) = p(\theta_1)[2y + (1-\rho)p(\theta_I)(w_I + q(\theta_R)y - y)] - (K+c)\theta_1, \tag{36}$$

where the arguments on the right-hand side are suppressed.

Now we have

$$\frac{d\hat{F}}{d\rho} = \frac{\partial \hat{F}}{\partial \theta_1}\frac{d\theta_1}{d\rho} + \frac{\partial \hat{F}}{\partial \theta_I}\frac{d\theta_I}{d\rho} + \frac{\partial \hat{F}}{\partial w_I}\frac{dw_I}{d\rho} + \frac{\partial \hat{F}}{\partial \rho}. \tag{37}$$

We will go through each term in turn. First, from lemma 1 we know that the first term is zero. Second,

$$
\begin{aligned}
\frac{\partial \hat{F}}{\partial \theta_I} &= p(\theta_1)\frac{d}{d\theta_I}\left\{(1-\rho)p(\theta_I)[w_I + q(\frac{(1-\rho)p_I(\theta_I)p_1}{1-p_1})y - y]\right\} \\
&= (1-\rho)p(\theta_1)[w_I + (1-\epsilon)q_R y - y]\frac{dp(\theta_I)}{d\theta_I} \\
&= (1-\rho)p(\theta_1)[w_I + V_R - y]\frac{dp(\theta_I)}{d\theta_I} > 0.
\end{aligned}
$$

To get from the first to the second equation we used the fact that $d(p_I q(\theta_R))/dp_I = (1-\epsilon)q_R$. To get from the second to the third equation we used that $(1-\epsilon)q_R y = V_R$. The inequality follows from the wage floor imposed on w_I. Since we are investigating the case where θ_I is strictly decreasing in ρ (if not we know welfare is falling in ρ), it follows that the second term in (37) is strictly negative.

From (36) it follows that $\partial \hat{F}/\partial w_I = p(\theta_1)(1-\rho)p(\theta_I) > 0$. In appendix 10.10 we show that $dw_I/d\rho < 0$. Hence the third term in (37) is also negative. Finally, we have that

$$
\begin{aligned}
\frac{\partial \hat{F}}{\partial \rho} &= p(\theta_1)\frac{d}{d\rho}\left\{(1-\rho)p(\theta_I)[w_I + q(\frac{(1-\rho)p_I(\theta_I)p_1}{1-p_1})y - y]\right\} \\
&= -p(\theta_1)p(\theta_I)[w_I + V_R - y] < 0,
\end{aligned}
$$

where the steps are similar to the steps for $\partial \hat{F}/\partial \theta_I$. Hence all the terms in (37) are negative or zero, and we have shown the following proposition:

Proposition 8 *Covenants not to compete reduce welfare.*

Covenants not to compete lower welfare even if they reduce the probability of losing a worker to imitation firms. To get more intuition, first note that since $w_I \geq y - V_R$, the surplus a worker creates in an imitation firm is at least as large as that from a worker that stays in the innovation firm. Thus, the presence of imitation firms in itself is good for welfare. However, imitation would be more valuable if fewer workers would leave the innovation firm, and those who leave get a higher wage w_I. This is, in effect, what happens when the imitation firm pays a mandatory transfer to the incumbent firm, or, when entry of imitation firms is taxed (although in the latter case it is the government, and not the workers, that collects the increase in surplus). A covenant not to compete, in contrast, destroys resources, which means it reduces the matching rate without giving higher wages in return. Therefore, welfare decreases.

Next we analyze the innovating firm's incentive to implement such covenants not to compete. Consider first a scenario where firms can commit to ρ at the hiring stage in period 1. In appendix

10.11 we show that M_1 (and thus M_2) is strictly decreasing in ρ. Hence, not surprisingly, firms will always find it in their interest to set $\rho = 0$.

Consider then a scenario where firms set ρ at the beginning of period 2. An employee can avoid the constraint on her job search by quitting before search takes place. Hence, analogous to the situation with a quit fee, the contract the firm offers has to satisfy the interim participation constraint of the worker. More specifically, the firm offers a contract $\{\rho, w_2\}$ that satisfies $W_2 \geq \bar{W}$, or

$$\overline{W} \leq (1-\rho)p(\theta_I)w_I + (1-(1-\rho)p(\theta_I))w_2 \tag{38}$$

and, in addition, $w_2 \geq U_2$. As above, this latter inequality is always satisfied when (38) is satisfied.

An issue is how the workers' and the firms' search behavior is influenced by the covenant not to compete. The problem is a reformulation of equation (33):

$$\{\widehat{\theta}_I, \widehat{w}_I\} = \underset{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I=0}{\arg\max} \ [w_2 + (1-\rho)p(\tilde{\theta}_I)(\tilde{w}_I \ - w_2)].$$

Suppose imitation firms cannot observe individual firms' choice of ρ. Then the constraint $V_I = 0$ is independent of a single firm's choice of ρ, and it follows that $\bar{\theta}_I$ and $\bar{w}_I$ are independent of ρ ($1-\rho$ is just a multiplier). One can show that if the imitation firms observe ρ, this has the same effect on $\{\bar{\theta}_I, \bar{w}_I\}$ as scaling up c to $c/(1-\rho)$, in which case $\bar{\theta}_I$ and $\bar{w}_I$ both fall. In what follows we assume the latter, although our results holds in both cases.

Suppose first that w_2^{**} is equal to U_2, where w_2^{**} denotes the limited commitment wage in the absence of covenants not to compete. In this case (38) binds, hence the firm has to compensate the worker if $\rho > 0$. Further, as in the quit fee case, when (38) binds, the objective function of the innovating firm can be written as $M_2 - \overline{W}$. Then, since M_2 is strictly decreasing in ρ, the firm sets $\rho = 0$.

Suppose next that $w_2^{**} > U_2$. In this case the constraint (38) does not bind at $\rho = 0$, and the firm may set $\rho > 0$ without increasing w_2. Since the firm's period-2 profit J_2 is decreasing in ρ (see appendix 10.11 for a formal proof) it is in the firm's interest to set $\rho > 0$. The argument applies up to the value $\bar{\rho}$ at which (38) starts to bind. Hence, the firm will set $\rho = \bar{\rho}$. We have the following proposition:

Proposition 9 *Suppose w_2^{**}, the limited-commitment wage in the absence of covenants not to* *compete, strictly exceeds U_2. Then the innovation firms set $\rho > 0$.*

The intuition for the result is that if $w_2^{**} > U_2$, the worker receives a rent by staying on in the firm. Hence, if the firm increases ρ slightly, it can do this without compensating the worker, it only dissipates some of this rent. As a result the firm has an incentive to increase ρ up to the point at which the outside option of the worker binds.

Hence innovation firms may have an incentive to set ρ strictly higher than zero in some situations. However, we know from above that covenants not to compete always reduce welfare. Our analysis thus clearly indicates that courts should not enforce covenants not to compete clauses.

Renting out Workers Finally, consider the scenario where the innovating firm has the possibility of renting out the worker to an imitation firm. In this case it is the innovating firm that does the search for a job, and it faces the same frictions as the worker does when she would search on the job. Since the firm has all the bargaining power, the interim participation constraint will again bind, and the worker receives an expected income of $\overline{W}$ as defined in (32). Denote the rental price to the imitating firm as w_I^r. The innovating firm in period 2 now maximizes:

$$\max_{\widetilde{\theta}_I, \widetilde{w}_I^r \ s.\ to\ V_I=0} p(\widetilde{\theta}_I)(V_R + \widetilde{w}_I^r) + (1 - p(\widetilde{\theta}_I))y - \overline{W}$$

The maximand can be written as $y - \overline{W} + p(\theta_I)(w_I^r - (y - V_R))$. Since $y - V_R$ is equal to the full-commitment wage w_2, the firm's problem is equivalent to the worker's maximization in the full commitment case up to a constant. It follows that the solution is efficient.

Proposition 10 *When the innovating firms can rent out workers to imitating firms, the equilibrium allocation $\{\theta_1, \theta_I\}$ is efficient.*

The intuition is straightforward. The firm is residual claimant on the value of search, and hence searches efficiently.

8 Discussion

Our model builds on several seemingly strong assumptions. In this section we will discuss some of them in more detail.

As mentioned in the description of the model environment, an important assumption is that a single firm cannot expand indefinitely, but can hire at most one worker. As argued below, a maximum capacity of one worker can be thought of as a normalization, the important assumption is that firms are small relative to the market. Hence there is room for many firms that pay the innovation cost and make a profit from the innovation. As in many models of monopolistic competition, the scarce factor of production is labor,[29] and firms enter the market up to the point where the tightness of the labor market makes innovation just worthwhile. The most direct interpretation of limited firm size is technological, i.e., that the production function of each firm exhibits decreasing returns to scale. Limited firm size may also be interpreted as a reduced form model of product differentiation, as in the standard Dixit-Stiglitz framework. With this interpretation, each innovator creates a new product variety, and aggregate demand for each product is limited.

Further, we could allow for multi-worker firms as in Pissarides (2000) and Kaas and Kircher (2011)), as long as the firms are small relative to the market and hence act as price takers. Suppose each innovation firm hires up to n workers, and that the output is proportional to the number of employees up to the capacity limit. For each position, the firm opens one vacancy, which is filled with a probability $p(\theta_1)$. Suppose also that all workers in an innovating firm learn about the innovation. Finally, suppose that the innovation cost is nK. It is then straightforward to show that this model is isomorphic to our model, with the same equilibrium characteristics and welfare properties. In particular, the policy recommendations will still hold. Likewise, our model can also easily be extended to allow for an expansion of innovating firms, for instance by allowing innovating firms to hire one more worker from the replacement market in period 2. This allows the innovating firms to exploit the non-rivalry of the knowledge use in-house. In all other respects, the model is as before, in particular the incumbent worker does on-the-job search. Technically, the new element of the model is that innovating firms post two vacancies in the replacement market if the incumbent worker has moved on, and one if the incumbent worker stays, instead of one and zero as in the original version. Everything else equal, this will increase the tightness in the replacement market and hence drive up w_2, both with full and limited commitment. This will tend to reduce the amount of entry by imitating firms. The effect on the amount of entry of innovation firms is not clear. On the one hand, the hiring opportunity is also a profit opportunity. This will tend to increase entry.

[29] See for instance Melitz (2003).

On the other hand, the increased tightness in the replacement market will reduce period 2 profits. In addition, the outside option of available workers in period 1 (which is to enter the replacement market and cash in U_2 in period 2) will increase. The latter two effects go in the direction of a reduced entry. Hence the net effect is unclear. More important, however, is that exactly the same externalities will be present as in the original model. With full commitment, the imitation search market will maximize the profit of the incumbent worker and firm pair, without creating externalities. Hence the equilibrium will be efficient. With limited commitment, the period-2 wages paid by innovation firms to workers with knowledge will be too low to deliver efficiency, and too many imitation firms will enter the market. Hence the inefficiencies analyzed in the original model prevail. Our conjecture is that our policy results also hold with this extension.

Our assumption that an entrepreneur who innovates has to attract a worker from the pool of available workers is not a crucial one. We can easily adjust our model so that innovators have a worker readily available without costs, because she is already hired. The entrepreneur offers her employee a contract that satisfies the worker's participation constraint. We do not expect any qualitative changes in the outcomes from this modification. First, the search stage in period 1 in the original model is not the source of any inefficiencies. Second, the key element of our model, i.e. the search market in period 2, still remains in place.

As already mentioned, a key element of our model is that the set of available workers is limited. This is crucial for our policy results. For instance, our result that a subsidy to innovation increases welfare depends on this assumption, as the effect goes through the labor market tightness in the replacement market. The set of available workers may be interpreted as the set of workers already working in the relevant sector or industry. The crucial element of our model is that the supply of labor above some point is upward-sloping in wages, in which case, say, more vacancies in the replacement market will drive up wages.

Finally, our model is set in two periods. Extending our model to one with an infinite horizon is on our agenda for future work. This would allow us to analyze the dynamic effects of policies and would make our framework more comparable to the related models in the growth literature.

9 Conclusion

In this paper we propose a model of innovation, imitation and spillovers through worker flows, in which the worker flows are explicitly modelled by using the Diamond-Mortensen-Pissarides matching framework with wage posting. We analyze under what circumstances the decentralized equilibrium of the model gives rise to an efficient allocation of resources. We find that the equilibrium is efficient if innovation firms can commit to long-term wage contracts with their workers. In the limited-commitment case, in which such contracts are absent, there is too little innovation and a too high probability of hiring by imitation firms in equilibrium compared with the efficient solution.

Our model allows us to analyze the effects of various policies, as well as the welfare effects of firm-level measures aimed at reducing turnover. In the limited commitment case we find that subsidizing innovation and taxing imitation improves welfare. Firm level measures like quit fees as well as innovation firms renting out workers to imitation firms also improve welfare. By contrast, covenants not to compete, interpreted as destruction of matches between innovation and imitation firms, always reduce welfare.

References

Acemoglu, D. K., and J.-S. Pischke (1999): "The structure of wages and investment in general training," *Journal of Political Economy*, 107(3), 539–572.

Aghion, P., and P. Howitt (1992): "A model of growth through creative destruction," *Econometrica*, 60(2), 323–351.

Akcigit, U., M. A. Celik, and J. Greenwood (2013): "Buy, keep or sell: Theory and evidence from patent resales," Working Paper 19763, National Bureau of Economic Research.

Almeida, P., and B. Kogut (1999): "Localization of knowledge and the mobility of engineers in regional networks," *Management Science*, 45(7), 905–917.

Arrow, K. (1962): "Economic welfare and the allocation of resources for invention," in *The rate and direction of inventive activity: Economic and social factors*, pp. 609–626. NBER.

Balsvik, R. (2011): "Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing," *The Review of Economics and Statistics*, 93(1), 285–297.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying technology spillovers and product market rivalry," *Econometrica*, 81(4), 1347–1393.

Boeri, T., P. Garibaldi, and E. R. Moen (2013): "The economics of severance pay," IZA Discussion Paper 7455, Institute for the Study of Labor (IZA).

Breschi, S., and F. Lissoni (2009): "Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows," *Journal of Economic Geography*, 9(4), 439–468.

Chiu, J., C. Meh, and R. Wright (2011): "Innovation and growth with financial, and other, frictions," Discussion paper, National Bureau of Economic Research.

Combes, P.-P., and G. Duranton (2006): "Labour pooling, labour poaching, and spatial clustering," *Regional Science and Urban Economics*, 36(1), 1–28.

Cooper, D. P. (2001): "Innovation and reciprocal externalities: information transmission via job mobility," *Journal of Economic Behavior & Organization*, 45(4), 403–425.

Davidson, C., and P. Segerstrom (1998): "R&D subsidies and economic growth," *The RAND Journal of Economics*, 29(3), 548–577.

Diamond, P. A. (1982): "Wage determination and efficiency in search equilibrium," *The Review of Economic Studies*, 49(2), 217–227.

Eeckhout, J., and B. Jovanovic (2002): "Knowledge spillovers and inequality," *The American Economic Review*, 92(5), 1290–1307.

Fernández-Blanco, J. (2013): "Labor market equilibrium with rehiring," *International Economic Review*, 54(3), 885–914.

Fosfuri, A., and T. Rønde (2004): "High-tech clusters, technology spillovers, and trade secret laws," *International Journal of Industrial Organization*, 22(1), 45–65.

Görg, H., and E. Strobl (2005): "Spillovers from foreign firms through worker mobility: An empirical investigation," *The Scandinavian Journal of Economics*, 107(4), 693–709.

Greenstone, M., R. Hornbeck, and E. Moretti (2010): "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of Political Economy*, 118(3), 536–598.

Grossman, G. M., and E. Helpman (1993): *Innovation and growth in the global economy*, vol. 1. The MIT Press.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic localization of knowledge spillovers as evidenced by patent citations," *The Quarterly Journal of Economics*, 108(3), 577–598.

Jones, C. I. (2005): "Growth and ideas," in *Handbook of Economic Growth*, vol. 1, pp. 1063–1111. Elsevier.

Kaas, L., and P. Kircher (2011): "Efficient firm dynamics in a frictional labor market," IZA Discussion Papers 5452, Institute for the Study of Labor (IZA).

Kim, J., and G. Marschke (2005): "Labor mobility of scientists, technological diffusion, and the firm's patenting decision," *RAND Journal of Economics*, 36(2), 298–317.

König, M., J. Lorenz, and F. Zilibotti (2012): "Innovation vs imitation and the evolution of productivity distributions," CEPR Discussion Papers 8843, C.E.P.R. Discussion Papers.

Lucas, R. E., and B. Moll (2011): "Knowledge growth and the allocation of time," Working Paper 17495, National Bureau of Economic Research.

Marimon, R., and V. Quadrini (2011): "Competition, human capital and income inequality with limited commitment," *Journal of Economic Theory*, 146(3), 976–1008.

Melitz, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 71(6), 1695–1725.

Menzio, G., and S. Shi (2010): "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, 145(4), 1453–1494.

Moen, E. R. (1997): "Competitive search equilibrium," *Journal of Political Economy*, 105(2), 385–411.

Moen, E. R., and Å. Rosén (2004): "Does poaching distort training?," *The Review of Economic Studies*, 71(4), 1143–1162.

Møen, J. (2005): "Is mobility of technical personnel a source of R&D spillovers?," *Journal of Labor Economics*, 23(1), 81–114.

Mortensen, D. T. (1982): "The matching process as a noncooperative bargaining game," in *The economics of information and uncertainty*, pp. 233–258. University of Chicago Press.

Mukoyama, T. (2003): "Innovation, imitation, and growth with cumulative technology," *Journal of Monetary Economics*, 50(2), 361–380.

OECD (2013): "Main science and technology indicators," http://www.oecd.org/sti/msti.

Pakes, A., and S. Nitzan (1983): "Optimum contracts for research personnel, research employment, and the establishment of "rival" enterprises," *Journal of Labor Economics*, 1(4), 345–65.

Pesola, H. (2011): "Labour mobility and returns to experience in foreign firms," *The Scandinavian Journal of Economics*, 113(3), 637–664.

Pissarides, C. A. (1985): "Short-run equilibrium dynamics of unemployment vacancies, and real wages," *American Economic Review*, 75(4), 676–90.

——— (2000): *Equilibrium unemployment theory*. MIT press.

Poole, J. P. (2013): "Knowledge transfers from multinational to domestic firms: evidence from worker mobility," *Review of Economics and Statistics*, 95(2), 393–406.

Romer, P. M. (1990): "Endogenous technological change," *Journal of Political Economy*, 98(5), S71–S102.

Rudanko, L. (2009): "Labor market dynamics under long-term wage contracting," *Journal of Monetary Economics*, 56(2), 170–183.

Segerstrom, P. S. (1991): "Innovation, imitation, and economic growth," *Journal of Political Economy*, 99(4), 807–827.

——— (2007): "Intel economics," *International Economic Review*, 48(1), 247–280.

Shi, S. (2009): "Directed search for equilibrium wage–tenure contracts," *Econometrica*, 77(2), 561–584.

Silveira, R., and R. Wright (2010): "Search and the market for ideas," *Journal of Economic Theory*, 145(4), 1550–1573.

Stoyanov, A., and N. Zubanov (2012): "Productivity spillovers across firms through worker mobility," *American Economic Journal: Applied Economics*, 4(2), 168–198.

The Economist (2013): "Schumpeter - Ties that bind," December 14th, 63.

10 Appendix

10.1 Deriving w_I

The imitating firm's problem is

$$\max_{\tilde{\theta}_I, \tilde{w}_I} \; q(\tilde{\theta}_I)(y - \tilde{w}_I) - c$$
$$\text{s. to} \quad W_2 \leq p(\tilde{\theta}_I)\tilde{w}_I + (1 - p(\tilde{\theta}_I))w_2.$$

Eliminating w_I by substitution from the constraint at equality, and using $p(\theta_I) = \theta_I q(\theta_I)$, the problem reads

$$\max_{\tilde{\theta}_I} \; \{q(\tilde{\theta}_I)(y - w_2) + \frac{w_2 - W_2}{\tilde{\theta}_I} - c\}.$$

The first order condition writes

$$q'(\theta_I)(y - w_2) - \frac{w_2 - W_2}{(\theta_I)^2} = 0.$$

Using $q'(\theta_I) = -\frac{\epsilon q(\theta_I)}{\theta_I}$ and substituting for W_2 using the constraint, the first order condition can be written

$$w_I = \epsilon y + (1 - \epsilon)w_2.$$

10.2 Deriving the optimal w_2

First we determine the derivative $\frac{d\widehat{p}_I(w_2)}{dw_2}$. Using $p_I(\theta_I) \equiv \theta_I q(\theta_I)$ together with equation (14), the zero-profit condition of imitating firms can be written

$$\widehat{p}_I(w_2) = \frac{\widehat{\theta}_I(w_2)c}{(1 - \epsilon)(y - w_2)}.$$

Totally differentiating this yields

$$d\widehat{p}_I(w_2) = \frac{c}{(1-\epsilon)(y-w_2)}\left(\frac{\widehat{\theta}_I(w_2)}{y-w_2}dw_2 + d\widehat{\theta}_I(w_2)\right) = \frac{\widehat{p}_I(w_2)}{y-w_2}dw_2 + q(\widehat{\theta}_I(w_2))d\widehat{\theta}_I(w_2),$$

where the second equality uses equation (14) once again. Then note that $dp(\theta_I) = d(\theta_I q(\theta_I)) = q(\theta_I)(1 + \frac{dq(\theta_I)}{d\theta_I}\frac{\theta_I}{q(\theta_I)})d\theta_I = q(\theta_I)(1-\epsilon)d\theta_I$. Therefore we can reformulate the previous expression to

$$d\widehat{p}_I(w_2) = \frac{\widehat{p}_I(w_2)}{y-w_2}dw_2 + \frac{1}{(1-\epsilon)}d\widehat{p}_I(w_2),$$

which finally yields

$$\frac{d\widehat{p}_I(w_2)}{dw_2} = -\frac{1-\epsilon}{\epsilon}\frac{\widehat{p}_I(w_2)}{y-w_2} \leq 0, \tag{39}$$

where $\frac{d\widehat{p}_I(w_2)}{dw_2} = 0$ if $w_2 \geq y$, since it follows from (14) that $\widehat{p}_I(w_2) = 0$ if $w_2 > y - c/(1-\epsilon)$.

In the main text we use (39) in (20), together with the envelope theorem, to obtain our result for w_2. Though, we can also achieve this result by plugging in w_I from (13) into (20) to get

$$\begin{aligned}\frac{dJ_1}{dw_2} &= \frac{d}{dw_2}[\widehat{p}_I(w_2)[V_R - (y - \epsilon y - (1-\epsilon)w_2)]] \\ &= \frac{d\widehat{p}_I(w_2)}{dw_2}[V_R - (1-\epsilon)(y-w_2)] + \widehat{p}_I(w_2)(1-\epsilon),\end{aligned}$$

and then substitute out $\frac{d\widehat{p}_I(w_2)}{dw_2}$ from (39), setting the expression equal to zero, and finally get

$$w_2 = y - V_R.$$

Next, we establish sufficiency of the first order condition at the solution and optimality with respect to the corner solution where $\widehat{p}_I(w_2) = 0$. First, the second derivative of $\widehat{p}_I(w_2)$ is

$$\frac{d^2\widehat{p}_I(w_2)}{(dw_2)^2} = -\frac{1-\epsilon}{\epsilon^2}(2\epsilon - 1)\frac{\widehat{p}_I(w_2)}{(y-w_2)^2}. \tag{40}$$

With this and the expression (39) it follows that

$$\begin{aligned}\frac{d^2 J_2}{(dw_2)^2} &= \frac{d^2\widehat{p}_I(w_2)}{(dw_2)^2}[V_R + (1-\epsilon)(w_2 - y)] + 2\frac{d\widehat{p}_I(w_2)}{dw_2} < 0 \\ &\Leftrightarrow -(2\epsilon - 1)\epsilon(y-w_2) - 2\epsilon(y-w_2) < 0,\end{aligned}$$

where we have used that at the optimum $V_R = y - w_2$.

Finally, we rule out that the innovation firm wants to set w_2 so high that it is not profitable for imitation firms to enter, i.e. $\widehat{p}_I(w_2) = 0$. We have to compare joint income in this case (which is $M_2 = y$) to joint income where $\widehat{p}_I(w_2)$ is positive:

$$\begin{aligned} y + \widehat{p}_I(w_2)(V_R + w_I - y) &> y \\ V_R + w_I - y &> 0 \\ \epsilon V_R &> 0, \end{aligned}$$

where in the last line, by using (13), the interior solution, $V_R = y - w_2$, has been substituted in.

10.3 Proof proposition 1

We first establish that there is a unique equilibrium in period 2 for given θ_1, and that the θ_I solving the second period equilibrium is strictly decreasing in θ_1.

Lemma 2 *For given θ_1 there is a unique $\theta_I(\theta_1)$ that satisfies the zero-profit condition (24). Furthermore, $\theta_I(\theta_1)$ is strictly decreasing.*

Proof. Define the value function $V_I(\theta_I;\theta_1)$ as

$$V_I(\theta_I;\theta_1) \equiv q(\theta_I)q(\theta_R)y(1-\epsilon)^2 - c.$$

We have that

$$\frac{\partial V_I(\theta_I;\theta_1)}{\partial \theta_I} = q'(\theta_I)q(\theta_R)y(1-\epsilon)^2 + q'(\theta_R)q(\theta_I)y(1-\epsilon)^2\frac{d\theta_R}{d\theta_I} < 0,$$

which directly follows from the definitions of the matching functions and the definition of θ_R. Furthermore, by assumption we have $V_I(0;\theta_1) = (1-\epsilon)^2 y - c > 0$, while $\lim_{\theta_I\to\infty} V_I(\theta_I;\theta_1) = -c$. Thus the equation has a unique solution for given θ_1. Finally, $V_I(\theta_I;\theta_1)$ is strictly decreasing in θ_1 by the definition of θ_R. Hence $\theta_I(\theta_1)$ is a strictly decreasing function. ■

To show existence and uniqueness of the overall equilibrium define

$$V_1(\theta_1) \equiv q(\theta_1)(1-\epsilon)[2y + \max_{\tilde{\theta}_I,\tilde{w}_I \text{ s. to } V_I=0}\{p(\tilde{\theta}_I)[V_R + \tilde{w}_I - y]\} - U_2] - K - c. \tag{41}$$

First note that $V_1(0) \geq (1-\epsilon)2y - c - K > 0$ by assumption, while $\lim\limits_{\theta_1\to\infty} V_1(\theta_1) = -c - K$, hence the equation has at last one solution.

To show that there is a unique θ_1 that solves this equation, it is sufficient to show that $\frac{\partial V_1(\theta_1)}{\partial \theta_1} < 0$. To this end, it follows from (41) that it is sufficient to show that

$$\frac{d}{d\theta_1}[2y + \max_{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I=0} \{p(\tilde{\theta}_I)[V_R + \tilde{w}_I - y]\} - U_2] \leq 0.$$

Using the envelope theorem, we find that

$$\frac{d}{d\theta_1}[2y + \max_{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I=0} \{p(\tilde{\theta}_I)[V_R + \tilde{w}_I - y]\} - U_2]$$
$$= p(\theta_I)\frac{dV_R}{d\theta_1} - \frac{dU_2}{d\theta_1}.$$

It thus suffices to show that $\theta_R = \frac{p(\theta_1)p(\theta_I(\theta_1))}{1-p(\theta_1)}$ is increasing in θ_1. Suppose not, i.e., suppose θ_I falls so much that θ_R decreases. Then V_R increases so that $w_2 = y - V_R$ must decrease. But then it follows from the zero-profit condition for the imitating firms (14) that θ_I increases, a contradiction. This together with lemma 2 proves the result.

10.4 Details of the Efficiency Result

In the following we provide the main steps of how to arrive at the first order conditions given in the text. The first order condition with respect to θ_I can be written as

$$\frac{\partial F}{\partial \theta_I} = p(\theta_1)y[p'(\theta_I)q(\theta_R) + p(\theta_I)\frac{dq(\theta_R)}{d\theta_I}] - p(\theta_1)c = 0,$$

where, by definition of the matching function, $p'(\theta_I) = q(\theta_I)(1-\epsilon)$, and

$$\frac{dq(\theta_R)}{d\theta_I} = -\frac{q'(\theta_R)}{\theta_R^2}\frac{d\theta_R}{d\theta_I} = -\frac{\epsilon q(\theta_R)p(\theta_1)q(\theta_I)(1-\epsilon)}{\theta_R(1-p(\theta_1))} = -\frac{\epsilon q(\theta_R)q(\theta_I)(1-\epsilon)}{p(\theta_I)}.$$

The first order condition can then be rearranged to

$$p(\theta_1)[q(\theta_I)q(\theta_R)y(1-\epsilon)^2 - c] = 0. \tag{42}$$

The first order condition with respect to θ_1 reads

$$\frac{\partial F}{\partial \theta_1} = y\{p'(\theta_1)[2 + p(\theta_I)q(\theta_R) - \frac{c\theta_I}{y}] + p(\theta_1)p(\theta_I)\frac{dq(\theta_R)}{d\theta_1} - \frac{c+K}{y}\} = 0,$$

where $p'(\theta_1) = q(\theta_1)(1-\epsilon)$, and $\frac{dq(\theta_R)}{d\theta_1} = -\frac{q'(\theta_R)}{\theta_R^2}\frac{d\theta_R}{d\theta_1} = -\frac{\epsilon q(\theta_R)p(\theta_I)(1-\epsilon)q(\theta_1)}{\theta_R(1-p(\theta_1))^2} = -\frac{\epsilon q(\theta_R)}{p(\theta_1)}\frac{(1-\epsilon)q(\theta_1)}{(1-p(\theta_1))}$.
The first order condition can then be rearranged to:

$$q(\theta_1)(1-\epsilon)[2 + p(\theta_I)q(\theta_R) - c\theta_I] - p(\theta_1)[\frac{p(\theta_I)\epsilon(1-\epsilon)q(\theta_R)q(\theta_1)}{p(\theta_1)(1-p(\theta_1))}] - \frac{c+K}{y} = 0$$
$$\Longleftrightarrow$$
$$q(\theta_1)y(1-\epsilon)[2 + p(\theta_I)q(\theta_R)\epsilon(1-\epsilon) - \epsilon p(\theta_R)] - c - K = 0,$$

where we have used $c = yq(\theta_I)q(\theta_R)(1-\epsilon)^2$ from (42).

10.5 Deriving the posting wage w_{2A}

First we derive the wage w_{2A} and then we show that this wage gives a local maximum. The first order condition of the firm's period-2 problem given by (28) reads

$$\frac{dJ_2}{dw_{2A}} = p(\theta_I) - 1 + \frac{d\widehat{p}_I(w_2)}{dw_{2A}}[V_R + w_{2A} - y] = 0. \tag{43}$$

Substituting in the expression for $\frac{d\widehat{p}_I(w_2)}{dw_{2A}}$ from (39) and solving for w_{2A} gives

$$w_{2A} = \left(y - \frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon}V_R\right), \tag{44}$$

Note that in cases where $p(\theta_I) \leq \epsilon$ in equilibrium, the marginal increase in profits when marginally raising w_{2A} is weakly negative. Thus, the firm will optimally set w_{2A} equal to the lower bound, U_2, in this case. Incorporating the lower bound, w_{2A} is then given by:

$$w_{2A} = \begin{cases} max\{(y - \frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon}V_R), U_2\}, & \text{if } p(\theta_I) > \epsilon; \\ U_2, & \text{otherwise.} \end{cases} \tag{45}$$

Finally, we establish sufficiency at the interior solution for w_{2A}. We need to show that

$$\frac{d^2J_2}{(dw_{2A})^2} = \frac{d^2\widehat{p}_I(w_2)}{(dw_{2A})^2}[V_R + w_{2A} - y] + 2\frac{d\widehat{p}_I(w_2)}{dw_{2A}} < 0.$$

Using the expression for $\frac{d\widehat{p}_I(w_2)}{dw_{2A}}$ from (39) and the second derivative analogous to (40):

$$\frac{d^2\widehat{p}_I(w_2)}{(dw_{2A})^2} = -\frac{1-\epsilon}{\epsilon^2}(2\epsilon - 1)\frac{\widehat{p}_I(w_2)}{(y - w_{2A})^2},$$

we get

$$\begin{aligned} \frac{d^2J_2}{(dw_{2A})^2} &= -\frac{1-\epsilon}{\epsilon}\frac{\widehat{p}_I(w_2)}{y - w_{2A}}\left[(\frac{2\epsilon-1}{\epsilon})(\frac{V_R}{y-w_{2A}} - 1) + 2\right] < 0 \\ &\Leftrightarrow (\frac{2\epsilon-1}{\epsilon})(\frac{V_R}{y-w_{2A}} - 1) + 2 > 0 \\ &\Leftrightarrow \frac{1}{\epsilon}\left[(2\epsilon-1)\frac{V_R}{y-w_{2A}} + 1\right] > 0 \\ &\Leftarrow \ w_{2A} < y - V_R. \end{aligned}$$

The second to the last line follows from the fact that $2\epsilon - 1 > -1$. The last line is implied by (44) since $\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon} > 1$. This proves the result.

10.6 Proof of proposition 3 and of proposition 4

We first establish the following two lemmata:

Lemma 3 *For given θ_1, define $\theta_I(\theta_1)$ as the solution to the zero-profit condition of the imitating firms given optimal wages (29) or (30). $\theta_I(\theta_1)$ is a decreasing single-valued function.*

Proof. The optimal wage and the zero-profit condition is given by:

$$w_I = w_2 + \epsilon(y - w_2)$$
$$q(\theta_I) = \frac{c}{(1-\epsilon)(y - w_2)}, \tag{46}$$

when w_I is not at the wage floor constraint. The left-hand side of (46) is strictly decreasing in θ_I while the right-hand side is increasing in θ_I (since θ_R is increasing in θ_I). By assumption, $c < (1-\epsilon)^2 y$, thus the left-hand side is strictly greater than the right-hand side for $\theta_I = 0$. This ensures existence and uniqueness. An increase in θ_1 shifts θ_R up and therefore the right-hand side of (46) increases, hence θ_I falls. When the imitating firm's wage is bound by the wage floor constraint, i.e. $w_I = \breve{w}_I \equiv y - V_R$, the zero-profit condition is given by

$$q(\theta_I) = \frac{c}{(y - \breve{w}_I)}. \tag{47}$$

Note that $\breve{w}_I$ is increasing in θ_R (since V_R is decreasing in θ_R). Thus, the result that θ_I falls is shown by a similar ument as before. ■

Lemma 4 *For given θ_1, θ_I is strictly higher and w_2 strictly lower than in the full commitment case.*

Proof. In the following the uments are based on the equilibrium outcome of the second period for a given entry of firms in period 1. We will denote equilibrium values for given θ_1 of a variable x as $x^*(\theta_1)$ and $x^{**}(\theta_1)$ for the full commitment and limited commitment case, respectively. Now, by contradiction suppose the opposite of the lemma is true, i.e. $\theta_I^*(\theta_1) \geq \theta_I^{**}(\theta_1)$. Then $\theta_R^*(\theta_1) \geq \theta_R^{**}(\theta_1)$ and hence $V_R^*(\theta_1) \leq V_R^{**}(\theta_1)$. Thus, $w_2^*(\theta_1) = y - V_R^*(\theta_1) \geq y - V_R^{**}(\theta_1) > w_2^{**}(\theta_1)$, and hence by (46) $\theta_I^*(\theta_1) < \theta_I^{**}(\theta_1)$, a contradiction. Further, given $\theta_I^*(\theta_1) < \theta_I^{**}(\theta_1)$, condition (46) implies $w_2^{**}(\theta_1) < w_2^*(\theta_1)$.

When w_I is given by the wage floor constraint, i.e. $w_I = \check{w}_I \equiv y - V_R$, the zero-profit condition is given by

$$q(\theta_I) = \frac{c}{(y - w_I)}. \tag{48}$$

To show the result in this case, again suppose the opposite is true, i.e. $\theta_I^*(\theta_1) \geq \theta_I^{**}(\theta_1)$. Then $\theta_R^*(\theta_1) \geq \theta_R^{**}(\theta_1)$ and hence $V_R^*(\theta_1) \leq V_R^{**}(\theta_1)$. Thus, $w_I^*(\theta_1) = w_2^*(\theta_1) + \epsilon(y - w_2^*(\theta_1)) = y - (1 - \epsilon)V_R^*(\theta_1) > y - V_R^{**}(\theta_1) \equiv \check{w}_I(\theta_1)$, and hence by (48) $\theta_I^*(\theta_1) < \theta_I^{**}(\theta_1)$, a contradiction. ■

Turning now to the main proof, we follow the notation introduced in lemma 4. The value of the innovating firm in the first period can then be written as a function of θ_1:[30]

$$\begin{aligned} V_1^{**}(\theta_1) &= q(\theta_1)(1-\epsilon)[2y + p(\theta_I)[V_R + w_I - y] - U_2] - K - c \qquad (49) \\ &= q(\theta_1)(1-\epsilon)[2y + p(\theta_I^{**}(\theta_1))[q(\theta_R)(1-\epsilon)y + (1-\epsilon)(w_2^{**}(\theta_1) - y)] - p(\theta_R)\epsilon y] - K - c, \end{aligned}$$

for the case in which the wage floor constraint for w_I does not bind. We have to show that a solution to $V_1^{**}(\theta_1) = 0$ exists. First, note that since w_2 for either wage setting case is continuous in θ_1, $V_1^{**}(\theta_1)$ is also continuous. Also, by assumption we have that $V_1^{**}(0) \geq (1-\epsilon)2y - K - c > 0$. Furthermore $\lim_{\theta_1 \to \infty} V_1^{**}(\theta_1) = -K - c < 0$. Hence, it follows from the intermediate value theorem that an equilibrium exists.

To prove proposition 4, insert for U_2 in equation (41) for the full commitment case to get

$$V_1^*(\theta_1) = q(\theta_1)(1-\epsilon)[2y + \max_{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I = 0} \{p(\tilde{\theta}_I)[V_R + \tilde{w}_I - y]\} - p(\theta_R)\epsilon y] - K - c.$$

For given θ_1 and V_R, the term within the max operator of $V_1^*(\theta_1)$ compares to the corresponding term of $V_1^{**}(\theta_1)$ in the following way:

$$\max_{\tilde{\theta}_I, \tilde{w}_I \ s.\ to\ V_I = 0} \{p(\tilde{\theta}_I)[V_R + \tilde{w}_I - y]\} \geq p(\theta_I^{**}(\theta_1))[V_R + w_I^{**}(\theta_1) - y].$$

Furthermore, we know from lemma 4 that for given θ_1, $\theta_I^{**}(\theta_1) > \theta_I^*(\theta_1)$. Hence $\theta_R^{**}(\theta_1) > \theta_R^*(\theta_1)$ and therefore V_R ($p(\theta_R)$) is higher (lower) in the full commitment case for given θ_1. By termwise comparison it then follows that $V_1^*(\theta_1) > V_1^{**}(\theta_1)$. Since $V_1^*(\theta_1)$ is strictly decreasing in θ_1, as established in proposition 1, it follows that $\theta_1^* > \theta_1^{**}$.

[30] For convenience we suppress the dependence of θ_R on θ_1.

When w_I is at the wage floor, i.e. $w_I = \check{w}_I \equiv y - V_R$ we can write the profits of the innovating firms in equilibrium as:

$$V_1^{**}(\theta_1) = q(\theta_1)(1-\epsilon)[2y + p(\theta_I)[V_R + \check{w}_I - y] - p(\theta_R)\epsilon y] - K - c$$
$$= q(\theta_1)(1-\epsilon)[2y - p(\theta_R)\epsilon y] - K - c.$$

Following a similar ument as above, showing existence is straightforward. To show proposition 4 when $w_I = \check{w}_I$, first note that in the full commitment case $\max_{\tilde{\theta}_I, \tilde{w}_I \ |V_I=0} \{p(\tilde{\theta}_I)[V_R + \tilde{w}_I - y]\} > 0$. Next it follows from lemma 4 that for given θ_1, $p(\theta_R)$ is lower in the full commitment case. Hence, we have the result $V_1^*(\theta_1) > V_1^{**}(\theta_1)$. By the same ument as above, we can conclude that $\theta_1^* > \theta_1^{**}$.

10.7 Proof of lemma 1

First we show that $\frac{\partial F(\theta_1^{**}, \theta_I^{**})}{\partial \theta_1} = 0$. To this end, rewrite the welfare function:

$$F(\theta_1, \theta_I) = p(\theta_1)2y + (1 - p(\theta_1))p(\theta_R)y - \theta_1(c + K) - p(\theta_1)c\theta_I,$$

using $\theta_R = \frac{p(\theta_I)p(\theta_1)}{1-p(\theta_1)}$. Taking derivatives with respect to $\theta_1$we get

$$\frac{\partial F}{\partial \theta_1} = q(\theta_1)(1-\epsilon)[2y - c\theta_I] + \frac{d}{d\theta_1}[(1-p(\theta_1))p(\theta_R)y] - (c+K).$$

Remember that the free entry condition of imitating firms is given by

$$V_I = q(\theta_I)(y - w_I) - c = 0, \tag{50}$$

so we know that $c = q(\theta_I)(y - w_I)$ in equilibrium. Insert this into the derivative to get

$$\frac{\partial F}{\partial \theta_1} = q(\theta_1)(1-\epsilon)[2y + p(\theta_I)(w_I - y)] + \frac{d}{d\theta_1}[(1-p(\theta_1))p(\theta_R)y] - (c+K). \tag{51}$$

Next, note that $dp(\theta_i) = q(\theta_i)(1-\epsilon)d\theta_i$ for any market i. Then, again using $\theta_R = \frac{p(\theta_I)p(\theta_1)}{1-p(\theta_1)}$, we have

$$\frac{d}{dp(\theta_1)}[(1-p(\theta_1))p(\frac{p(\theta_I)p(\theta_1)}{1-p(\theta_1)})y]$$
$$= [-p(\theta_R) + (1-p(\theta_1))p'(\theta_R)\frac{p(\theta_I)p(\theta_1)}{(1-p(\theta_1))^2} + (1-p(\theta_1))p'(\theta_R)\frac{p(\theta_I)}{1-p(\theta_1)}]y$$
$$= -p(\theta_R)\epsilon y + p(\theta_I)q(\theta_R)(1-\epsilon)y$$
$$= -U_2 + p(\theta_I)V_R.$$

It then follows that

$$\frac{d}{d\theta_1}[(1-p(\theta_1))p(\theta_R)y] = q(\theta_1)(1-\epsilon)[-U_2 + p(\theta_I)V_R],$$

which inserted into (51) gives

$$\begin{aligned}\frac{\partial F}{\partial \theta_1} &= q(\theta_1)(1-\epsilon)[2y + p(\theta_I)(V_R + w_I - y) - U_2] - (c+K) \\ &= q(\theta_1)(1-\epsilon)[M_1 - U_2] - (c+K).\end{aligned}$$

Now we will compare this first order condition to the zero-profit condition in the limited commitment case. In effect, when the firm chooses w_1 in period 1, it takes M_1 as given and maximizes $V_1 = q(\theta_1)(M_1 - W_1) - c - K$ subject to (3), which gives the first order condition $W_1 = \epsilon M_1 + (1-\epsilon)U_2$. Using this to substitute out W_1 from V_1 gives

$$V_1 = q(\theta_1)(1-\epsilon)[M_1 - U_2] - (c+K).$$

Hence, the first order condition for the socially optimal θ_1 is equal to the zero-profit condition of innovating firms.

Next, we establish the second condition, $\frac{\partial F(\theta_1^{**},\theta_I^{**})}{\partial \theta_I} < 0$. Recall that the derivative of the welfare function with respect to θ_I is $\frac{\partial F}{\partial \theta_I} = p(\theta_1)[q(\theta_I)q(\theta_R)y(1-\epsilon)^2 - c]$. We will evaluate this derivative at the limited commitment allocation. Substitute out c from (50) to get

$$\frac{\partial F}{\partial \theta_I} = p(\theta_1)[q(\theta_I)q(\theta_R)y(1-\epsilon)^2 - q(\theta_I)(1-\epsilon)(y - w_2)] \ ,$$

where we have used that $w_I = \epsilon y + (1-\epsilon)w_2$. Then we substitute in for w_2 from our two wage setting mechanism to determine the sign of this derivative for any pair $[\theta_1\, , \theta_I]$. In the wage posting case we get

$$\frac{\partial F}{\partial \theta_I} = p(\theta_1)[q(\theta_I)q(\theta_R)y(1-\epsilon)^2 - q(\theta_I)q(\theta_R)y(1-\epsilon)^2(\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon})] < 0,$$

where the inequality follows from the fact that $\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon}$ is larger than 1 (when $w_2 > U_2$). In the wage bargaining case we get

$$\frac{\partial F}{\partial \theta_I} = p(\theta_1)[q(\theta_I)q(\theta_R)y(1-\epsilon)^2 - q(\theta_I)(1-\epsilon)(y - \beta(y - V_R) - (1-\beta)U_2)] < 0,$$

where the inequality follows from that since $U_2 < y - V_R$, we have $q(\theta_I)(1-\epsilon)(y - \beta(y-V_R) - (1-\beta)U_2) > q(\theta_I)(1-\epsilon)(y - \beta(y-V_R) - (1-\beta)(y-V_R)) = q(\theta_I)q(\theta_R)y(1-\epsilon)^2$. Hence, at the limited commitment equilibrium allocation, the derivative of the welfare function with respect to θ_I is negative (proofs for when w_2 is on the bound or w_I is on the wage floor follow exactly the same line of ument and are therefore omitted).

10.8 Proof of proposition 5

It is immediate that a subsidy shifts V_1 up and thus increases θ_1. Furthermore,

$$\frac{dF}{d\theta_1} = \frac{\partial F}{\partial \theta_1} + \frac{\partial F}{\partial \theta_I}\frac{d\theta_I}{d\theta_1} = \frac{\partial F}{\partial \theta_I}\frac{d\theta_I}{d\theta_1} > 0,$$

where $\frac{\partial F}{\partial \theta_1} = 0$ and $\frac{\partial F}{\partial \theta_I} < 0$ by lemma 1. Then, the inequality follows from the fact that higher θ_1 implies lower θ_I as stated in lemma 3.

10.9 Proof of proposition 6

As has been established in the proof of lemma 3, the left-hand side of the zero-profit condition for the imitating firms, $q_I(\theta_I) = c/(y - w_I)$, decreases with θ_I, whereas the right-hand side increases, regardless of whether w_I is interior or on the wage floor. Thus an increase in c through a tax decreases θ_I for a given θ_1. The induced effect of θ_1 through the other zero-profit condition could only overturn the decrease in θ_I, if θ_1 decreases sufficiently enough. By contradiction, assume that θ_I increases with the tax.[31] Then by the zero-profit condition of the imitating firms, w_{2A} must decrease. Since θ_I increases, θ_1 has to decrease sufficiently to lower θ_R in order for w_{2A} (as given in 29) to go down. Recall the zero-profit condition of the innovators:

$$q(\theta_1)(1-\epsilon)(M_1 - U_2) - K + c = 0,$$

where

$$M_1 = 2y + p(\theta_I)V_R(1 - \frac{p(\theta_I)(1-\epsilon)^2}{p(\theta_I) - \epsilon}).$$

Since θ_I increases and θ_R decreases, M_1 increases when $w_I + V_R > y$ (since $-\frac{p_I(1-\epsilon)^2}{p_I - \epsilon}$ is increasing in θ_I, and V_R is decreasing in θ_R). Furthermore, U_2 decreases. Thus, to satisfy the zero-profit condition of innovating firms, θ_1 has to increase, a contradiction. It follows that θ_1 cannot decrease that much, hence a tax reduces reduces θ_I. The case of $w_I = y - V_R$ can be established in a similar way. Welfare then increases due to lemma 1.

[31] The following is for wage setting case A only. Case B can be shown along the same lines and is therefore omitted.

10.10 Proof of derivative of w_I w.r.t. ρ

We ue that $\frac{dw_I}{d\rho} < 0$. By contradiction assume $\frac{dw_I}{d\rho} \geq 0$. We consider two cases: First, assume that θ_I increases with ρ in equilibrium. Then it follows immediately from the zero-profit condition of the imitators that w_I has to fall. Second, if θ_I decreases with ρ consider wage setting case A (case B follows exactly the same line of ument is therefore omitted). It follows from the equilibrium value of w_{2A} that w_{2A} (and thereby w_I) can increase if and only if θ_R increases. Since θ_I decreases, θ_1 has to increase sufficiently. Recall the zero-profit condition of the innovators:

$$q(\theta_1)(1-\epsilon)(M_1 - U_2) = K + c, \tag{52}$$

where

$$M_1 = 2y + (1-\rho)p(\theta_I)(w_I + V_R - y).$$

Suppose θ_1 increases so much that θ_R increases enough so that w_{2A} stays constant. Then it follows that M_1 falls if $w_I + V_R > y$. Furthermore, U_2 increases. Hence the left hand side of (52) decreases. Given that the equilibrium is locally stable, it follows that θ_1 cannot increase that much, hence the result follows. The case of $w_I = y - V_R$ can be established in a similar way.

10.11 Proof of derivative of M_1 with respect to ρ

Using the definitions of J_2 and W_2 we can write

$$M_1 = J_2 + W_2 + y.$$

Next by the envelope theorem, keeping θ_R fixed since we look at the firms problem, we have $\frac{dJ_2}{d\rho} = -p(\theta_I)[V_R + w_2 - y]$ and $\frac{dW_2}{d\rho} = \frac{dw_2}{d\rho}(1 + (1-\rho)p(\theta_I)) - p(\theta_I)(w_I - w_2)$ (note that $\frac{dJ_2}{d\rho} > 0$ since $w_2 < y - V_R$). Combining gives

$$\frac{dM_1}{d\rho} = -p(\theta_I)[V_R + w_I - y] + \frac{dw_2}{d\rho}(1 + (1-\rho)p(\theta_I)),$$

where the first part is negative due to the wage floor constraint, $w_I \geq y - V_R$. Then we obtain the result if $\frac{dw_2}{d\rho} \leq 0$. Let $w_2(\rho)$ be the solution to the innovating firm's period 2 maximization problem (28) with first order condition (derived analogously as in appendix 10.2):

$$w_2(\rho) = y - \frac{(1-\rho)p(\theta_I)(1-\epsilon)}{(1-\rho)p(\theta_I) - \epsilon} V_R, U_2\}.$$

To show the result, suppose the opposite is true, i.e. $\frac{dw_2}{d\rho} > 0$. Since $\frac{\partial w_2(\rho)}{\partial \rho} \leq 0$, $\frac{dw_2}{d\rho}$ can only be positive if $\frac{dp(\theta_I)}{d\rho} > 0$. From the imitation firm's zero-profit condition, $(1-\rho)q(\theta_I)(1-\epsilon)(y-w_2) = c$, it follows that a lower $q(\theta_I)$ together with a higher ρ implies a lower w_2. A contradiction.