Productivity Spillovers through Labor Mobility in Search Equilibrium*

Tom-Reiel Heggedal[†], Espen R. Moen[‡], Edgar Preugschat[§]

September 10, 2015

Abstract

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 innovating 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 innovating firms to charge quit fees or rent out workers to imitating firms also improves welfare. By contrast, non-pecuniary measures like restrictions on mobility, interpreted as reducing matching efficiency between imitating firms and workers from innovating firms, always reduce welfare.

Key words: Efficiency, innovation, imitation, productivity, search frictions, spillovers, worker flows.

JEL Codes: J63, J68, 031, 038.

*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 SAET conference in Paris 2013, for helpful comments and fruitful discussions. Thanks to the Norwegian Research Council for financial support.

[†]Norwegian Business School, Oslo, Norway, Email: tom-reiel.heggedal@bi.no

[‡]Norwegian Business School, Oslo, Email: espen.r.moen@bi.no (corresponding author)

[§]Technical University of Dortmund, Germany, Email: edgar.preugschat@tu-dortmund.de

1 Introduction

Productivity spillovers associated with R&D are considered to be important.[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. The recent empirical literature has identified labor mobility as an important channel for such spillovers.[2] If a worker moves from a technologically advanced firm to one that is less so, she may bring valuable knowledge with her.[3] Hence worker flows create information flows.

In this paper we analyze under what conditions information flows caused by labor flows give rise to positive externalities from innovation, and when they do, we identify which policy remedies can improve the allocation of resources. To this end we set up a two-period directed search model with on-the job search. Firms may enter as an innovating firm in period 1, or as an imitating firm in period 2. An innovating firm shares its productive idea with its worker and an imitating firm gains access to this knowledge if it hires such a worker. An innovating firm that loses a worker still possesses the required knowledge, and can therefore hire a new worker and continue production. However, as there are search frictions in the labor market, losing the worker is costly.

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 contracting environment for 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 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 trade off a higher rent by lowering the wage in the second period against a lower chance of retaining the worker. This leads to lower joint surplus in period 2 which is anticipated in period 1, implying less entry of innovating firms. On the other hand, imitation –by hiring workers from innovating firms– becomes cheaper, implying excessive entry of imitating firms. Hence, there is too little innovation and too much imitation in equilibrium compared with the social optimal levels. Without search frictions, the equilibrium allocation is efficient even in the absence of long-term contracts, as competition for workers with knowledge protects their long-term wages. Hence, it is the combination of search frictions and limited commitment that creates the inefficiency.

[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] We discuss the empirical literature in more detail below.

[3] This knowledge may for instance be intangible organizational capital transferred by mangers, see e.g. Lustig, Syverson, and Van Nieuwerburgh (2011) and Eisfeldt and Papanikolaou (2013).

We derive a simple and useful characterization of the social welfare function at the limited-commitment allocation, which allows us to evaluate various policy measures. 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.[4] This may seem trivial, however, the results crucially hinge on general equilibrium effects between labor markets.

Importantly, we also study the welfare implications of 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 covenants not to compete.[5] To what extent courts honour such contracts varies. For instance, due to different legal traditions, some states in the US enforce covenants not to compete clauses in employment contracts, whereas others are more reluctant to do so (see Gilson (1999)). The study by Saxenian (1996) suggests worker mobility as an important channel for interfirm knowledge transfers. She contrasts the high employee turnover region of Silicon Valley, where covenants not to compete are illegal, with the region of Route 128 on the East coast, where such clauses are enforced.

We model different aspects of real-world mobility restrictions to analyze the effects of each channel in isolation. We find that allowing innovating firms to charge quit fees or renting out workers to imitating firms improves efficiency. By contrast, restrictions on worker mobility, interpreted as reducing matching efficiency, always harm welfare. Still, firms may have an incentive to impose restrictions on mobility in order to reduce worker turnover and extract rents from workers ex post. Hence, it follows from our analysis that courts should be reluctant to enforce such contracts.

Spillovers as we model it have similarities with general training. In both cases the worker acquires knowledge at one firm which can be utilized by other firms the worker moves to. The difference is that with human capital investments (Becker (1962)), the investment is lost if the worker quits. With spillovers, the investing firm still has the knowledge, and the cost associated with the worker quitting is the replacement cost of the worker. The latter is endogenously determined in search equilibrium. It is this endogenous replacement cost that is the main channel for welfare improving policies in our paper. This replacement cost is absent in models of general human capital investments, e.g., Acemoglu (1997), Acemoglu and Pischke (1999) and Moen and Rosén (2004). The difference will be discussed in more detail below.

Related Literature. There are several strands of literature that relate to our work. First, spillovers are at the core of endogenous growth models with innovation and imitation.[6] Several papers, following the seminal work by Segerstrom (1991), also analyze optimal policy.[7] However, in these papers it is imposed by assumption that spillover effects through imitation give rise to positive

[4] 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)).

[5] 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.

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

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

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 labor 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.[8] In these papers the dissemination of ideas might be inefficient as innovating firms have an incentive to limit worker flows in order to prevent increased price competition in the product market. In our paper we abstract from product market competition and focus on the cost of information flows coming from the frictional hiring process. Such search frictions are essential, as without them equilibrium always reaches efficiency. To our knowledge, none of the papers in the industrial organization literature on imitation contains search frictions.

While our paper connects on a technical level to the literature on search with contracting under limited commitment,[9] we are not aware of any work that analyzes innovation and imitation within a labor-search environment.[10]

As noted above, our model is related to models with on-the-job investments in general human capital in the presence of search frictions. In Acemoglu (1997), there is suboptimal investments in training due to a hold-up problem. Workers and their new employer bargain over the terms of trade, and at that point in time the costs of the investments are sunk. Hence the poaching firm gets part of the gain from the investments. In our paper search is directed, and poaching firms compete for workers *ex ante*. There is no underlying hold-up problem in our model. The different effects of imitation and human capital investment on optimal policy can be seen most directly by comparing our paper with Moen and Rosén (2004), who study human capital investments with directed search and provide some policy analysis. In Moen and Rosén, the investment level in human capital is below its first best level. Still it is constrained efficient; a training subsidy would reduce welfare. In our model, by contrast, a subsidy on innovation improves welfare. Increased entry of innovating firms makes the replacement market more crowded, increases wages for workers with knowledge in innovating firms, and reduces entry of imitating firms. Interestingly, we can replicate the constrained-efficiency results of Moen and Rosén in our model if we assume that the innovating firm is without value if the worker quits.[11]

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

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

[10]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 (2014). Relatedly, Marimon and Quadrini (2011) study human capital accumulation on-the-job in a setting with limited commitment, but without search frictions.

[11]See the discussion section for further details.

Empirical Motivation. There is a substantial empirical literature that provides direct and indirect evidence on spillovers through worker flows.[12] 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. Section 2 describes the economy, while sections 3 and 4 analyze the equilibrium when firms can and can not commit to long-term wage contracts, respectively. Next, section 5 establishes efficiency of the equilibrium with full commitment and the inefficiency of the equilibrium with limited commitment. Then, in section 6, we discuss public policies (taxes and subsidies), while a detailed analysis of firm policies (quit fees, restrictions on mobility, and options of renting out workers) is undertaken in section 7. Section 8 provides a discussion of the differences between spillovers and human capital as well as some of our model assumptions. The last section concludes.

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.

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

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 innovating 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_1 in period 1 and – if the match is continued – produce y_2 in period 2. 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 firm can keep the innovation and can post a vacancy in the next period.

In the beginning of period 2, there is also entry of entrepreneurs. Instead of innovating themselves, these entrepreneurs can create imitating firms, and attempt to hire a worker from an innovating firm to learn the innovation from her. This is the source of spillovers in our model. For convenience, we assume that imitating firms incur the same costs for posting vacancies (c) as innovating firms and successful imitating firms produce $y_I \leq y_2$. Further, innovating firms that have lost their worker, still possess the relevant knowledge, and can produce $y_R \leq y_2$ with the help of a worker without the knowledge. The extent of spillovers in our model can vary across two dimensions: First, the transferability of the knowledge is captured by y_I as a fraction of y_2. This ratio might be below 1 if e.g. the technology is to some extent specific to the innovating firm, or if the worker cannot fully transfer the knowledge to the imitating firm. Second, the non-rivalrousness of the knowledge is captured by y_R as a fraction of y_2. A reason for imperfect non-rivalrousness could be for instance that some part of the knowledge is embodied in the worker who takes the knowledge with her. We refer to the case where $y_2 = y_I = y_R$ as perfect spillovers.

In period 2, both firms that were not matched in the first period and innovating firms that have lost their worker search to find an employee at cost c_R. These firms search among the workers that were not hired by a firm in period 1. We call this the replacement market. All innovating firms that are matched in period 2 produce y_R.

In period 2 entrepreneurs could enter to innovate at cost K. However, they will choose not to do so. Conceptually, we model the introduction of new technologies in two phases, which correspond directly to the two time periods. First, in the innovation phase, imitation is difficult and firms only innovate. When workers with knowledge become abundant, the economy moves to the imitation phase, in which firms find it more profitable to imitate. Technically, we ensure this outcome of no entry to innovation in period 2 by our assumption of no discounting.[13] However, there are alternative parameter restrictions to obtain the same result in the presence of discounting.[14]

Our way of modeling innovations is general enough to encompass a number of interpretations. Innovations may be new technologies, 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 most of our examples involve non-patented ideas, our analysis is also relevant for models with spillovers where ideas can be patented. Hiring workers from firms with a patented idea can lead to spillovers in the

[13]Profits of entering in period 2 net of innovation costs are the same as the period-2 profits of innovating firms that where unmatched in the first period. Thus, entering in the first period dominates entering in the second period.

[14]A sufficient parameter restriction would be $K + c > y_2$.

sense that the receiving firm can develop new ideas that build upon the patent.

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 a limited demand for each product. Note, however, that for the welfare analysis, it is important that a firm's productivity reflects both the social and the private value of production.

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 innovating firms cannot expand in period 2. In the discussion section we also study the effects of letting the innovating firms better exploit their own innovation by allowing them to expand by hiring an additional (but 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 1, 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.[15] Let $\theta_i \equiv v_i/s_i$ 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 $p(\theta_i) = \theta_i q(\theta_i)$. To simplify the notation, we use the shorthand $q_i \equiv q(\theta_i)$ and $p_i \equiv p(\theta_i)$ throughout the main text unless the explicit version is needed for clarity.

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 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.

[15]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.

The following summarizes the timing protocol:

First Period:

1. Entrepreneurs enter and pay cost K in order to innovate and create an innovating firm.
2. Each innovating 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_1 units of output, while unmatched firms keep their innovation but stay idle. Employed workers learn the innovation.

Second Period:

1. New entrepreneurs enter and set up an imitating firm at no costs.
2. The imitating firms post a vacancy for informed workers at cost c.
3. Innovating firms that have lost their worker as well as innovating firms that remained unmatched in the first period post a vacancy for the remaining available workers at cost c_R.
4. Matched firms produce: y_2 in continuing matches; y_R in innovating firms that are being matched in period 2; y_I in imitating firms. Other firms exit.

We focus on sets of parameter constellations such that output is large relative to the search costs so that all three labor markets operate. In particular, we require that c is small relative to y_I, c_R is small relative to y_R. In addition we require that K is not too small relative to y_2 so that the market is not over-crowded with innovating firms in period 1, and the bounds on the matching function described in footnote 15 do not bind. As the parameter set is not easily characterized we refer the reader to the existence proofs in the appendix for the details.

In the benchmark model we assume that firms can commit to long-term wage contracts. We then relax this assumption in section 4 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 innovating firm in the second period can break up the match before the imitation market opens, or quit to accept an offer from an imitating 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 innovating firms introduce restrictions on workers who move to imitating firms, for instance in the form of a quit fee. We therefore ignore this constraint until we study restrictions on turnover in section 7.

3 Model with Full Commitment

In this section we first set up the model and then analyze equilibrium when firms can commit to long-term wage contracts.

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

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

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_I is the probability of finding a job at an imitating firm.[16] 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 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

$$U_1 = p_1 W_1 + (1 - p_1) U_2 \tag{3}$$

and

$$U_2 = \ p_R w_R, \tag{4}$$

respectively, where w_R is the wage offered in the replacement market, and p_1 and p_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

$$J_1 = y_1 - w_1 + p_I V_R + (1 - p_I)(y_2 - w_2), \tag{5}$$

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

$$V_R = q_R(y_R - w_R) - c_R, \tag{6}$$

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

$$\theta_R = \frac{p_1 p_I + \theta_1(1 - q_1)}{1 - p_1}, \tag{7}$$

where the first summand of the numerator is derived from the fact that the measure of workers at 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 imitating firm. The second summand is the number of innovating firms that remain unmatched in the first period. The denominator gives the mass of searching workers, which is equal to the mass of workers that have not found a job in the first period.

[16] If $w_I \leq w_2$, workers will not search, and $p_I = 0$. Recall that the dependency of p_I on θ_I is suppressed.S

The ex-ante value of innovating and opening a vacancy in an innovating firm is

$$V_1 = q_1 J_1 + (1 - q_1) V_R - c - K, \tag{8}$$

where q_1 is the probability that the vacancy is filled. The value of a vacancy in an imitating firm is

$$V_I = q_I (y_I - w_I) - c, \tag{9}$$

where q_I is the job-filling probability.

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. In addition to the standard assumptions regarding advertised wages and the probability of hiring workers, innovating 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_I that the worker quits. We follow here the literature on competitive on-the-job search (see Moen and Rosén (2004), Shi (2009), and Menzio and Shi (2010)). Suppose a small subset of innovating firms offer a wage 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 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 $\hat{\theta}_I(w_2)$ and $\hat{w}_I(w_2)$, are given by[17]

$$\{\hat{\theta}_I(w_2), \hat{w}_I(w_2)\} = \underset{\theta_I, w_I \ s.\ to\ V_I = 0}{\arg\max} \ p_I w_I + (1 - p_I) w_2. \tag{10}$$

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

$$V_1 = q_1[y_1 - w_1 + \hat{p}_I(w_2) V_R + (1 - \hat{p}_I(w_2))(y_2 - w_2)] + (1 - q_1) V_R - c - K, \tag{11}$$

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

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

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]

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

This is the standard result in competitive search models: the surplus (here $y_I - w_2$) 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 $\hat{w}_I(w_2)$ in (9), the zero-profit condition for the imitating firms implicitly determines $\hat{\theta}_I(w_2)$:

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

Given the solution for $\hat{\theta}_I(w_2)$, we obtain $\hat{p}_I(w_2)$.

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

$$w_R = \epsilon y_R,$$

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:

$$V_R = q_R(1-\epsilon) y_R - c_R$$
$$U_2 = p_R \epsilon y_R.$$

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

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 $\hat{p}_I(w_2)$ and $\hat{w}_I(w_2)$.
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

$$M_2 = y_2 + \hat{p}_I(w_2)[V_R + \hat{w}_I(w_2) - y_2]. \tag{15}$$

The joint income of a matched worker-firm pair in period 1 is then given by

$$M_1 = y_1 + M_2. \tag{16}$$

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

$$\begin{aligned} J_1 = M_1 - W_1 &= y_1 + y_2 + \hat{p}_I(w_2)[V_R + \hat{w}_I(w_2) - y_2] - W_1 \\ &= y_1 + y_2 + \hat{p}_I(w_2)[V_R + w_2 - y_2] + \hat{p}_I(w_2)(\hat{w}_I(w_2) - w_2) - W_1. \end{aligned} \tag{17}$$

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

$$\begin{aligned} \frac{dJ_1}{dw_2} &= \frac{d\hat{p}_I(w_2)}{dw_2}[V_R + w_2 - y_2] + \hat{p}_I(w_2) + \frac{d}{dw_2}[\hat{p}_I(w_2)(\hat{w}_I(w_2) - w_2)] \\ &= 0. \end{aligned} \tag{18}$$

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

$$w_2 = y_2 - V_R.$$

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 imitating 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.

If we would reduce the degree of non-rivalry by reducing y_R, V_R would be lower and therefore the wage w_2 would move closer to y_2. This in turn reduces the incentives for imitation. In the

[19] See appendix 10.2 for more details.

extreme case we would have $V_R = 0$ and no imitation would occur.

Turning to the optimal recruiting problem in step two, the firm now takes M_1 as given and maximizes $V_1 = q_1(M_1 - W_1) + (1 - q_1)V_R - c - K$ subject to (3). The first-order condition

$$W_1 = \epsilon(M_1 - V_R) + (1 - \epsilon)U_2,$$

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

By substituting in equilibrium values into (8), the zero-profit condition for innovating firms can be written:

$$V_1 = q_1(1 - \epsilon)[M_1 - V_R - U_2] + V_R - c - K. \tag{19}$$

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

$$V_I = q_I(1 - \epsilon)[y_I - y_2 + V_R] - c = 0. \tag{20}$$

Here we see that if the transferability of technology were limited, i.e. $y_I < y_2$, imitation would be less profitable with no entry of imitating firms in the extreme case.

In appendix 10.3 we show the following result for the equilibrium allocation $\{\theta_1^*, \theta_I^*\}$:

Proposition 1 *For a range of parameter values, the equilibrium defined above exists and is unique.*

The proof, including restrictions on parameters, is given in appendix 10.3. The parameter restrictions ensure that the economy is sufficiently productive so that all the three markets, the innovation market is period 1, the imitation market in period 2, and the replacement market in period 2, are active. More specifically, the innovation must be sufficiently productive relative to the innovation cost so that firms will enter to innovate. In addition, the parameter values make it sufficiently attractive for innovating firms to enter the replacement market, which in turn makes room for imitating firms to enter.

4 Model with Limited Commitment

In the setting of the previous section firms can commit to future wages. This is arguably a strong assumption. Wage contracts which specify future wage growth are rarely seen in practice. In a world where asymmetric information make state-contingent contracts difficult to honor, binding long-term wage contracts may be costly. For instance, Boeri, Garibaldi, and Moen (2013) show that long-term wage contracts with high future wages - which cannot be made contingent on future productivity - may lead to excessive firing ex post. Firms may also want to have discretion over future wages to avoid opportunistic behavior and shirking by workers.

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 assume that innovating firm can match wage offers from imitating firms, similar to Postel-Vinay and Robin (2002). The wage paid by a successful imitating firm must then be at least $\bar{w}_I = y_2 - V_R$. Otherwise, a separation would be inefficient as the value of keeping the worker for the innovating firm is higher than searching for a new one. We refer to $\bar{w}_I$ as the lower bound on w_I.

The imitating firm's problem is then to choose the wage w_I so as to maximize V_I given by (9) subject to (2) and the new constraint $w_I \geq y_2 - V_R$. If the bound $\bar{w}_I$ does not bind, it follows that the wage is given by equation (13): $\hat{w}_I(w_2) = \epsilon y_I + (1-\epsilon)w_2$. The associated labor market tightness $\hat{\theta}_I(w_2)$ is implicitly defined by the zero profit condition for imitating firms, $q(\hat{\theta}_I(w_2))((1-\epsilon)(y_I - w_2) - c = 0$, and the probability that the worker leaves is $\hat{p}_I(w_2)$, as in the full commitment case. On the other hand, if w_2 is set so low that $\hat{w}_I(w_2) < \bar{w}_I$, i.e., if w_2 is set lower than $\bar{w}_2$ given by

$$\bar{w}_2 = \frac{y_2 - \epsilon y_I - V_R}{1-\epsilon},$$

then $w_I = \bar{w}_I$. If $w_2 < \bar{w}_2$, the imitating firm will still pay $\bar{w}_I$. Hence, for any $w_2 < \bar{w}_2$, it follows that $p_I = \hat{p}_I(\bar{w}_2) \equiv \bar{p}_I$.

The innovating firm posts a wage w_2 in period 2 before its worker's on-the-job search decision.[20] The firm's period-2 profit can be written as

$$J_2 = (1-p_I)(y_2 - w_2) + p_I V_R. \tag{21}$$

The participation constraint of worker requires that $w_2 \geq U_2$ and the innovation firm's maximization problem reads

$$\begin{aligned} &\max_{w_2} J_2 \\ \text{s. to} \quad & p_I = \begin{cases} \hat{p}_I(w_2) & \text{if } w_2 \geq \bar{w}_2 \\ \bar{p}_I & \text{if } w_2 < \bar{w}_2 \end{cases} \\ & w_2 \geq U_2. \end{aligned}$$

Suppose, neither the lower bound on w_2 nor the one on w_I binds. Then we show in appendix 10.4 that the first-order condition for w_2 is given by the equation

$$w_2 = y_I - \frac{\hat{p}_I(w_2)(1-\epsilon)}{\hat{p}_I(w_2) - \epsilon}(y_I - y_2 + V_R). \tag{22}$$

In the appendix we also show that this first-order condition always has a solution, and that the second order conditions are satisfied on the interval at which the bounds do not bind. Denote the solution to this equation by $\tilde{w}_2$.

For the general case, let w_2^{lc} be the profit-maximizing period-2 wage for an innovating firm. First, note that if $\tilde{w}_2 < U_2$, the lower bound on w_2 binds, hence $w_2^{lc} = U_2$. If $\tilde{w}_2 > U_2$, we have to check for the lower bound on w_I. If $\hat{w}_I(U_2) \geq \bar{w}_I$, we have that $w_2^{lc} = \tilde{w}_2$. If $\hat{w}_I(U_2) < \bar{w}_I$, we

[20] In an earlier version of the paper, (see Heggedal, Moen, and Preugschat (2014)), we show that our results are robust to assuming wage bargaining instead of wage posting of w_2.

have to compare the profit (given by (21)) for $w_2 = \tilde{w}_2$ and for $w_2 = U_2$, and then pick the wage that gives the higher value.[21] In all cases, w_2^{lc} can be written as a function of θ_R with the following properties:

Lemma 1 *The optimal period-2 wage can be expressed as a function $w_2^{lc} = w_2^{lc}(\theta_R)$ which is strictly increasing in θ_R. For a given θ_R, the period-2 wage in innovating firms is strictly lower in the limited commitment case than in the full commitment case.*

Proof. See appendix 10.4. ■

Taking θ_R as given, the lemma states that the second period wage w_2 is smaller in the limited commitment case. This is because the firm now trades off retention and rent extraction within the period. At the full commitment wage the firm is indifferent between keeping and losing the worker. By increasing second period profits when keeping the worker through lowering w_2, the firm can now increase overall profits.

Turning to period 1, innovating firms choose w_1 so as to maximize V_1 given by (11) subject to (3) and (12), with $w_2 = w_2^{lc}(\theta_R)$. In the appendix 10.5 we show the following proposition:

Proposition 2 *The limited commitment equilibrium exists under the same parameter restrictions as in the full-commitment case. If $w_2^{lc} = U_2$ (lower bound), the equilibrium is unique. If $w_2^{lc} = \hat{w}_2$ (interior wage), a sufficient (but not necessary) condition for uniqueness is that $\epsilon \leq 1/2$.*

Let us discuss the issue of uniqueness in some more detail. In order to show uniqueness, we must show that V_1 defined by (19) is decreasing in θ_1. At first glance this seems obvious: an increase in θ_1 makes both the labor market in period 1 and the replacement market in period 2 more crowded. This hurts innovating firms both in period 1 and in period 2 either if they do not find a worker in the first period or if they lost their worker to an imitating firm. However, a counteracting force is that w_I, the wage of a worker hired by an imitating firm, also increases, and this in isolation tends to increase the joint income of the worker and the firm. As long as the period-2 wage w_2 is equal to U_2, it is easy to see that this effect is too weak to offset the other effects that tend to reduce V_1. For the interior wage $\hat{w}_2$ this is not so easy to show. However, a sufficient, but not necessary parameter restriction for uniqueness is given by $\epsilon \leq 1/2$. Numerically, we have not been able to find cases of multiple equilibria for the parameter region where firms do not randomize between an interior solution and U_2 (see footnote 21). If they do randomize we cannot rule out that V_1, although continuous, is increasing in θ_1. In what follows we assume that the equilibrium is unique.[22]

We have shown above that w_2 is lower than in the full commitment case for any given θ_R. In fact, in appendix 10.5 we show that w_2 is lower than in the full commitment case for any level of

[21] A minor technical issue emerges here, as $w_2^{nc}(\theta_R)$ may be discontinuous at exactly one value of θ_R, and jump from U_2 to $\tilde{w}_2$. However, the value functions are still continuous, see appendix 10.5, particularly the proof of Lemma 3 for details.

[22] If it is not unique, we can make the refinement that we choose the one with the highest θ_1, which will always Pareto dominate other equilibria.

entry of innovating firms. This lower wage leads 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, θ_1 is also lower. This is established in the following proposition.

Proposition 3 *The limited-commitment equilibrium has a higher θ_I and a lower θ_1 than the full-commitment equilibrium.*

Proof. See appendix 10.6. ■

5 Efficiency

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

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 and constraints on information flows as the agents in the market. Since the mass of available workers is normalized to unity,[23] aggregate output in period 1 equals $p_1 y_1 - \theta_1(c+K)$. If a worker at an innovating firm moves to an imitating firm in period 2, her contribution to output is changed only by the difference between y_2 and y_I. However, the now vacant innovating firm will produce additional output only if it is able to hire a new worker. Aggregate net output therefore is

$$\begin{aligned} F(\theta_1, \theta_I) = & p_1[y_1 + y_2 + p_I(y_I - y_2) - c\theta_I] \\ & + (1-p_1)p_R y_R - (c+K)\theta_1 - c_R[p_1 p_I + \theta_1(1-q_1)], \end{aligned} \tag{23}$$

where, as before, θ_R is given by (7). The planner chooses θ_1 and θ_I so as to maximize welfare. Comparing the first-order conditions of the planner to the zero-profit conditions in full commitment equilibrium for innovators (19) and for imitators (20), we show in appendix 10.7 that they are indeed the same. Thus, the (necessary) equilibrium conditions are identical to the necessary conditions for the interior efficient allocation.

Proposition 4 *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 argument 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 imitating 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 innovating firm is exactly equal to the opportunity

[23] This implies both that p_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.

cost of letting the worker move to an imitating 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.

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.

We now turn to the limited commitment case. As shown in section 4 above, the limited commitment θ_I is higher and θ_1 is lower than the unique efficient allocation under full commitment. The following is immediate:

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

The intuition for the inefficiency result is as follows. First, with wage setting under limited commitment there is a downward pressure on the wage w_2. 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 (due to competitive search and 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 innovating 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_I. In period 1, the innovating firm may still extract the period 2 surplus from the worker, but the joint income is smaller than in the full commitment case. As a result, fewer innovating 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 (23). Let θ_1^{**} and θ_I^{**} denote the limited-commitment equilibrium values of θ_1 and θ_I, respectively. Then the following holds:

Lemma 2 *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 10.8. ■

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 provides a very helpful method for policy analysis. It implies that for any 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 Government Policies: 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. 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.[24]

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. Define $\sigma > 0$ as a marginal subsidy to vacancy creation in period 1 and $\tau > 0$ as a marginal tax on imitating firms' vacancy creation in period 2. We assume that any net receipts or losses are redistributed in a lump-sum fashion to all workers equally. Beginning with the subsidy, we have the following result:

Proposition 5 *A subsidy σ 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 10.9. ■

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 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.[25]

[24] 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.

[25] In section 8 we discuss this issue in the context of human capital.

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

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

Proof. See appendix 10.10. ■

A tax on imitating firms reduces the leaving rate of workers from innovating firms to imitating firms. This effect directly increases welfare.[26] Whether the 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 2 implies that this ambiguity is without consequence since $\partial F/\partial\theta_1 = 0$.

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. Firms may include covenants not to compete clauses in employment contracts to reduce turnover. Such clauses may take different forms. The milder form is that the worker has to compensate the firm if she leaves. We refer to this as a quit fee. A more drastic measure is that firms (or an employer association in an industry) introduce restrictions for workers on movements between firms. These restrictions on worker mobility can for instance prevent workers from forming matches with firms in an industry for a period of time. Below we analyze quit fees and restrictions on mobility separately to isolate the effects of different types of clauses in contracts. Finally, we also analyze the effects of firm options to rent out workers. Our analysis gives useful guidelines regarding the attitude the government should take towards these firm policies.

Quit Fees. 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 innovating 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.[27]

[26] In real life it may be hard to single out imitating firms and tax them. Note, however, that a tax on all firms and a subsidy on innovating firms will act as a tax on imitating firms. Note also that a high marginal tax rate on high incomes (paid by workers in imitating firms only) will reduce entry of imitating firms and in that sense act like a tax on such firms.

[27] It does so by influencing w_I through α. The ex-post payment to the worker when she pays a quit fee is $w_I = \epsilon y_I + (1-\epsilon)(w_2+\alpha)$. 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 of w_2.

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 a contract $\{w_2, \alpha\}$, which the worker accepts or rejects. This reduces the value of the worker of being employed in the innovating firm when searching for a job in an innovating firm. Recall that the worker may leave the firm at will at the beginning of period 2, before searching for a job in an imitating firm. The outside option of the worker at this stage reads

$$\overline{W} \equiv \max_{\theta_I, w_I \ s.\ to\ V_I=0} [p_I w_I + (1 - p_I) U_2]. \tag{24}$$

The associated participation constraint,

$$W_2 \geq \bar{W}, \tag{25}$$

until now always satisfied, may bind in this setting with quit fees. We refer to this constraint, somewhat imprecisely, as the interim participation constraint. This constraint implies that the firm has to offer a contract that gives at least $\bar{W}$. 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 innovating firm, is $W_2 = w_2 + p_I(w_I - \alpha - w_2)$. Analogous to (10), it follows that the values $\hat{\theta}_I$ and $\hat{w}_I$, as implicit functions of α and w_2, maximize W_2 given the zero profit constraint of imitating firms:

$$\{\hat{\theta}_I, \hat{w}_I\} = \arg\max_{\theta_I, w_I \ s.\ to\ V_I=0} [w_2 + p_I(w_I - \alpha - w_2)]. \tag{26}$$

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

$$J_2 = p(\hat{\theta}_I(w_2 + \alpha))(\alpha + V_R) + (1 - p(\hat{\theta}_I(w_2 + \alpha)))(y_2 - w_2), \tag{27}$$

with respect to w_2 and α, subject to (25).

The first thing to note is that the constraint (25) 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 $p(\hat{\theta}_I)$. However, the firm's ex-post profit would increase. Substituting (25) (which is binding so that $W_2 = \overline{W}$) into the expression for J_2 from above gives

$$\begin{aligned} J_2 &= p(\hat{\theta}_I(w_2 + \alpha))(\hat{w}_I + V_R) + (1 - p(\hat{\theta}_I(w_2 + \alpha)))y - \bar{W} \\ &= M_2 - \bar{W}, \end{aligned} \tag{28}$$

where M_2 is the joint income as defined in equation (15). This is similar 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 (28) with respect to w_2 is equivalent to the problem of maximizing M_1 given by (17) 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 $w_2 + \alpha = y_2 - V_R$.

To complete the analysis, insert the first order condition for w_I (analogous to (13), given by $w_I = \varepsilon y_I + (1-\varepsilon)(w_2+\alpha)$), and $w_2 + \alpha = y_2 - V_R$, into the expression for W_2 to obtain

$$\begin{aligned} W_2 &= w_2 + p(\hat{\theta}_I)(\epsilon y_I + (1-\epsilon)(y_2 - V_R) - (y_2 - V_R)) \\ &= w_2 + p(\hat{\theta}_I)\epsilon(y_I - y_2 + V_R). \end{aligned}$$

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 innovating firms will enter in period 1.

Restrictions on Mobility. 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 $(1-\rho)q_I$ (and the job finding probability in the imitation market by $(1-\rho)p_I$, where $\rho \in [0,1]$ is a measure of the strictness of the restrictions on mobility.

To understand the welfare effects of the restrictions on mobility, 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 (23) becomes

$$F(\theta_1, \theta_I, \rho) =$$
$$p_1[y_1 + y_2 + (1-\rho)p_I(y_I - y_2 + q_R y_R - c_R) - c\theta_I] + \theta_1(1-q_1)[q_R y_R - c_R] - (c+K)\theta_1.$$

It is straightforward to show that, for given values of θ_I and θ_1, an increase in ρ decreases welfare. However, an increase in ρ will also change the equilibrium values of θ_1 and θ_I. For a given ρ, the matching function is well defined, and Lemma 2 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, we know that 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 ($w_I \geq y_2 - V_R$).

It will prove to be convenient to substitute w_I back into the welfare function. Recall that in equilibrium $(1-\rho)q_I(y_I - w_I) = c$. Using this and the relationship $\theta_I q_I = p_I$, it follows that $c\theta_I = (1-\rho)p_I(y_I - w_I)$. If we substitute this into the expression for F, we find that the equilibrium welfare as a function of ρ writes

$$\hat{F}(\theta_1(\rho), \theta_I(\rho), w_I(\rho), \rho) = \tag{29}$$
$$p_1[y_1 + y_2 + (1-\rho)p_I(w_I + q_R y_R - y_2 - c_R)] + \theta_1(1-q_1)[q_R y_R - c_R] - (c+K)\theta_1.$$

Now we have

$$\frac{d\hat{F}}{d\rho} = \frac{\partial \hat{F}}{\partial \theta_1}\frac{d\theta_1}{d\rho} + \frac{\partial \hat{F}}{\partial \theta_I}\frac{d\theta_I}{d\rho} + \frac{\partial \hat{F}}{\partial w_I}\frac{dw_I}{d\rho} + \frac{\partial \hat{F}}{\partial \rho}. \tag{30}$$

In appendix 10.11 we show that each summand of (30) is either zero or negative, which gives the following proposition:

Proposition 8 *Restrictions on mobility reduce welfare.*

Restrictions on mobility lower welfare even if they reduce the probability of losing a worker to imitating firms. To get more intuition, first note that since $w_I \geq y_2 - V_R$, the surplus a worker creates in an imitating firm is at least as large as that from a worker that stays in the innovating firm. Thus, the presence of imitating firms in itself is good for welfare. However, imitation would be more valuable if fewer workers would leave the innovating firm, and those who leave get a higher wage w_I. This is, in effect, what happens when the moving worker pays a quit fee to the incumbent firm, or, when entry of imitating firms is taxed. A restriction on mobility, 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 restrictions on mobility. Consider first a scenario where firms can commit to ρ at the hiring stage in period 1. In appendix

10.12 we show that joint income, M_1, is strictly decreasing in ρ. Hence, 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

$$\overline{W} \leq (1-\rho)p_I w_I + (1-(1-\rho)p_I)w_2, \quad (31)$$

and, in addition, $w_2 \geq U_2$. As above, this latter inequality is always satisfied when (31) is satisfied.

An issue is how the workers' and the firms' search behavior is influenced by the restrictions on mobility. The problem is a reformulation of equation (26):

$$\{\hat{\theta}_I, \hat{w}_I\} = \underset{\theta_I, w_I \ s.\ to\ V_I=0}{\arg\max} \ [w_2 + (1-\rho)p_I(w_I - w_2)].$$

Suppose imitating 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 (i.e. $1-\rho$ is just a multiplier). One can show that if the imitating firms observe ρ, this has the same effect on $\{\bar{\theta}_I, \bar{w}_I\}$ as scaling up c of the imitating firm to $c/(1-\rho)$, in which case $\bar{\theta}_I$ and $\bar{w}_I$ both fall. In what follows we assume the latter scenario, although our result 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 restrictions on mobility. In this case (31) binds, hence the firm has to compensate the worker if $\rho > 0$. Further, as in the quit fee case, when (31) 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 (31) 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 increasing in ρ (see appendix 10.12 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 (31) 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 restrictions on mobility, strictly exceeds U_2. Then the innovating 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 appropriates 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 innovating firms may have an incentive to set ρ strictly higher than zero in some situations. However, we know from above that restrictions on mobility always reduce welfare. Our analysis thus clearly indicates that courts should not enforce covenants not to compete clauses that put restrictions on the mobility of workers.

Renting out Workers. Finally, consider the scenario where the innovating firm has the possibility of renting out the worker to an imitating 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 (24). Denote the rental price to the imitating firm as w_I^r. The innovating firm in period 2 now maximizes:

$$\max_{\theta_I, w_I^r \ s. \ to \ V_I=0} p_I(V_R + w_I^r) + (1 - p_I)y_2 - \overline{W}$$

The maximand can be written as $y_2 - \overline{W} + p_I(w_I^r - (y_2 - V_R))$. Since $y_2 - V_R$ 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 Extensions and Discussion

In this section we will discuss several of the assumptions of our model.

More on Human Capital vs Spillovers. Clearly, human capital and spillovers are related phenomena. However, when it comes to policy recommendations, there are some differences between the two that lead to different conclusions. To illustrate this, recall that the difference between spillovers and general human capital is captured by y_R, the productivity of the job if the worker leaves. With pure spillovers $y_R = y_2$, while with pure human capital y_R is low (zero).

Assume now that the productivity in the replacement market is so low that $V_R < 0$ for a tightness θ_R given by (7). Let $\bar{y}_R$ be such that, in equilibrium, $V_R(\theta_R; \bar{y}_R) = 0$, with θ_R given by (7). If y_R is below this threshold, then V_R is negative. Innovating firms with a vacancy will then randomize on whether to post the vacancy or not, and in the resulting mixed-strategy equilibrium, the tightness $\tilde{\theta}_R$ will be determined so that

$$V(\tilde{\theta}_R) = (1 - \epsilon)q((\tilde{\theta}_R)y_R - c_R = 0 \tag{32}$$

Equation (32) uniquely determines $\tilde{\theta}_R$ for $y_R \in [\frac{c_R}{1-\epsilon}, \bar{y}_R]$.[28]

In addition we relax the assumption that $y_I \leq y_2$, and assume instead that $y_I > y_2 + c$. With human capital investments, this may make sense; some firms (like research firms) have an advantage in training workers, others in utilizing the skills of trained workers. This ensures that we will still have entry of imitating firms.

The limited-commitment equilibrium of the model can be defined as above, but with (7) replaced by (32). We refer to this as the training equilibrium. Furthermore, it follows from the welfare properties of competitive search equilibrium that the mixed-strategy equilibrium in the replacement market is efficient, in the sense that it maximizes aggregate output less search costs. It is straightforward to show that Proposition 3 still holds in the training equilibrium, i.e., that the equilibrium allocation has less innovation and more imitation than the optimal allocation. Hence we can show the following proposition

Proposition 11 *In the training equilibrium, a subsidy σ on training vacancies reduces welfare.*

The result is analogous to the constrained-efficiency result in Moen and Rosén (2004). The proof is straightforward: The training subsidy affects θ_1. However, the wage w_2 does not change as θ_R is uniquely determined by (32). Therefore the maximization problem of the imitating firm, is unaltered by a training subsidy, and thus also θ_I. It follows that for a given θ_1, the period-2 profits of innovating firms are independent of the training subsidy. Hence the training subsidy unambiguously increases θ_1. However, we know from Lemma 2 that θ_1 is constrained optimal given θ_I with $\sigma = 0$, and the proposition follows.

In comparison, when there are spillovers, i.e. $y_R > \bar{y}_R$, the equilibrium is as in the main model and the effects of a subsidy to training go through the replacement market. A subsidy increases the number of training firms entering the market, the number of imitating firms entering (for a given tightness θ_I), and the tightness of the replacement market. As a result, the value of entering the replacement market for innovating firms that have lost their worker falls, and the innovating firms therefore protect their workers more. This reduces the incentives of imitating firms to enter the market, and it is this effect that improves welfare. In the training equilibrium, this effect is defused.

Observe that Proposition 11 also holds when $y_R = 0$, in which case the replacement market shuts down. Further note that a tax on imitating firms still improves welfare, as this has a direct effect on θ_I, independently of the replacement market.

The Role of Search Frictions. It may be enlightening to analyze the Walrasian equilibrium without search frictions and vacancy costs. To simplify we assume that $y_1 = y_2 = y_R = y_I = y$ and that $y < K < 2y$, and that search costs are zero (this is not necessary for our argument). The matching function can be written as $\min\{As_i^{\epsilon}v_i^{1-\epsilon}, s_i, v_i\}$ (see footnote 15). In the limit as A goes

[28]Recall from footnote 15. that the matching function can be written as $\min\{As_i^{\epsilon}v_i^{1-\epsilon}, s_i, v_i\}$. For sufficiently low values of θ_i, it follows that the number of matches is given by v_i. Here we assume that this is not the case.

to infinity, the matching function reduces to $\min\{s, v\}$. In equilibrium, $s = v$ in period 1, where s is the measure of workers that do search. At this point the elasticity of q with respect to θ is not well defined. However, the competitive equilibrium can easily be derived by observing that it must satisfy the following requirements:

1. The zero-profit constraints of both innovating firms in period 1 and imitating firms in period 2.

2. Workers at the beginning of period 1 are indifferent between getting a job in period 1 or waiting to get a job in the replacement market in period 2.

3. All workers are employed in period 2.

These requirements uniquely pin down the equilibrium where: (i) a measure 1/2 of innovating firms enter in period 1 and hire half of the work force, (ii) a measure 1/2 of imitating firms enter the market in period 2 and hire all employed workers, and (iii) the innovating firms hire all the remaining available workers in the replacement market.[29]

On average, a 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 $y \cdot 3/2 - K/2$. 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 innovating firms, as competition between imitating firms always increases the wage paid by imitating firms up to y. Hence, search frictions are key for our inefficiency results; without search frictions the equilibrium is efficient.

Limited Firm Size. 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,[30] 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 under monopolistic competition, 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. We conjecture that our welfare results will still

[29] 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 imitating firms, and w_R the wage paid by innovating firms to their new hires in period 2. Note, w_1 is negative since $K > y$.

[30] See for instance Melitz (2003).

hold if firms are allowed to price discriminate so that the social value of opening a market is equal to the private value to the firm.

The Diamond-Mortensen-Pissarides framework models a one-good economy. We conjecture that our analysis will still hold if we extend the model to allow for many goods, with downward sloping aggregate demand curves, as long as the individual firms are price takers. The important issue is that the private value and the social value of entering the market coincide. On the other hand, our analysis abstracts from strategic considerations that may arise if firms have market power. Intuitively, one would think that if firms have market power, and this leads to a deadweight loss, imitation may lead to more firms having access to the technology and thereby erode the market power of the innovating firms. This may increase the social value of imitation. Hence, our analysis is less relevant for markets in which firms have substantial market power and where this leads to deadweight losses.

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 innovating 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_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 innovating firms is not clear. On the one hand, the hiring opportunity is also a profit opportunity. This will tend to increase entry. 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 innovating firms to workers with knowledge will be too low to deliver efficiency, and too many imitating 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.

Timing of Innovation. 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.

Multiple Periods. For simplicity our model is set in two periods. We can extend the model so that within each period there are two stages (corresponding to the periods of the model in this paper), first an innovation stage and then an imitation stage. In this extended model the qualitative trade-offs for the firms are the same as in the two-period model, and it can be shown that an equilibrium with full commitment is efficient while an equilibrium with limited commitment is not.[31] Taking our framework to a to an infinite-horizon, endogenous-growth setting is on the agenda for future work. This will 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 mobility, 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 innovating 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 imitating firms in equilibrium compared with the efficient allocation.

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. Moreover, allowing innovating firms to charge quit fees or rent out workers to imitating firms also improves welfare. By contrast, restrictions on mobility, interpreted as reducing matching efficiency between innovating and imitating firms, always reduce welfare.

[31] A formal model of this extension is available upon request.

References

Acemoglu, D. (1997): "Training and innovation in an imperfect labour market," *The Review of Economic Studies*, 64(3), 445–464.

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.

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.

Becker, G. S. (1962): "Investment in human capital: A theoretical analysis," *The journal of political economy*, 70(5), 9–49.

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.

Eisfeldt, A. L., and D. Papanikolaou (2013): "Organization capital and the cross-section of expected returns," *The Journal of Finance*, 68(4), 1365–1406.

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.

Gilson, R. J. (1999): "Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete, The," *NYUl Rev.*, 74, 575.

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.

Heggedal, T.-R., E. R. Moen, and E. Preugschat (2014): "Productivity Spillovers Through Labor Mobility," .

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 (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122(1).

Lustig, H., C. Syverson, and S. Van Nieuwerburgh (2011): "Technological change and the growing inequality in managerial compensation," *Journal of Financial Economics*, 99(3), 601–627.

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.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 70(6), 2295–2350.

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.

SAXENIAN, A. (1996): *Regional advantage*. Harvard University Press.

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

$$\max_{\theta_I, w_I} \; q(\theta_I)(y_I - w_I) - c$$
$$\text{s. to} \quad W_2 \leq p(\theta_I)w_I + (1 - p(\theta_I))w_2.$$

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

$$\max_{\theta_I} \; \{q(\theta_I)(y_I - w_2) + \frac{w_2 - W_2}{\theta_I} - c\}.$$

The first-order condition writes

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

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

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

10.2 Deriving the Optimal w_2

First we determine the derivative $\frac{d\hat{p}_I(w_2)}{dw_2}$. 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

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

Totally differentiating this yields

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

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

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

which finally yields

$$\frac{d\hat{p}_I(w_2)}{dw_2} = -\frac{1-\epsilon}{\epsilon}\frac{\hat{p}_I(w_2)}{y_I - w_2} \leq 0, \tag{33}$$

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

In the main text we use (33) in (18), 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 (18) to get

$$\begin{aligned}\frac{dJ_1}{dw_2} &= \frac{d}{dw_2}[\hat{p}_I(w_2)[V_R - (y_2 - \epsilon y_I - (1-\epsilon)w_2)]] \\ &= \frac{d\hat{p}_I(w_2)}{dw_2}[V_R - y_2 + \epsilon y_I + (1-\epsilon)w_2)] + \hat{p}_I(w_2)(1-\epsilon),\end{aligned}$$

and then substitute out $\frac{d\hat{p}_I(w_2)}{dw_2}$ from (33), setting the expression equal to zero, and finally get

$$w_2 = y_2 - V_R.$$

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

$$\frac{d^2\hat{p}_I(w_2)}{(dw_2)^2} = -\frac{1-\epsilon}{\epsilon^2}(2\epsilon - 1)\frac{\hat{p}_I(w_2)}{(y_I - w_2)^2}. \tag{34}$$

With this and the expression (33) it follows that

$$\frac{d^2 J_2}{(dw_2)^2} = \frac{d^2\hat{p}_I(w_2)}{(dw_2)^2}[V_R + (1-\epsilon)(w_2 - y_2)] + 2\frac{d\hat{p}_I(w_2)}{dw_2} < 0$$
$$\Leftrightarrow -(2\epsilon - 1)\epsilon(y_2 - w_2) - 2\epsilon(y_I - w_2) < 0,$$

where we have used that at the optimum $V_R = y_2 - w_2$.

Finally, we rule out that the innovating firm wants to set w_2 so high that it is not profitable for imitating firms to enter, i.e. $\hat{p}_I(w_2) = 0$. We have to compare joint income in this case (which is $M_2 = y_2$) to joint income where $\hat{p}_I(w_2)$ is positive:

$$y + \hat{p}_I(w_2)(V_R + w_I - y_2) > y_2$$
$$V_R + w_I - y_2 > 0$$
$$\epsilon V_R > 0,$$

where in the last line, by using (13), the interior solution, $V_R = y_2 - w_2$, has been substituted in.

10.3 Proof of Proposition 1

We will show that an equilibrium in which all the three submarket operates exists and is unique under the following parameter restrictions:

$$(1-\epsilon)(y_I - y_2 + (1-\epsilon)y_R - c_R) - c > 0. \tag{35}$$

In addition, K must be neither too large nor too small. Note that (35) implies that the replacement value can be strictly positive.

We first establish that for given θ_1 there is a unique $\theta_I(\theta_1)$. Furthermore, $\theta_I(\theta_1)$ is weakly decreasing and continuous. To this end, define the function $V_I(\theta_I;\theta_1)$

$$V_I(\theta_I;\theta_1) \equiv q(\theta_I)(1-\epsilon)[y_I - y_2 + V_R] - c.$$

Let $\theta_1^1 > 0$ be defined by $V_I(0;\theta_1^1) = 0$. Note, that the value θ_1^1 is unique. First consider the case where $\theta_1 \leq \theta_1^1$.

When $\theta_1 \to 0$ we have that a $V_R \to (1-\epsilon)y_R - c_R > 0$ by assumption (with an interior solution for the matching function). Hence $V_I(0;\theta_1)$ goes to $(1-\epsilon)[y_I - y_2 + (1-\epsilon)y_R - c_R] - c$ which is strictly greater than zero by assumption. Further, $\lim_{\theta_I \to \infty} V_I(\theta_I;\theta_1) = -c$. As $V_I(\theta_I;\theta_1)$ is continuous, we have established that there exists a solution $\theta_I(\theta_1)$.

Furthermore, we have

$$\frac{\partial V_I(\theta_I;\theta_1)}{\partial \theta_I} = q'(\theta_I)[y_I - y_2 + V_R] + q(\theta_I)(1-\epsilon)\frac{\partial V_R}{\partial \theta_R}\frac{d\theta_R}{d\theta_I} < 0,$$

which follows from $\frac{\partial V_R}{\partial \theta_R} = (1-\epsilon)q'(\theta_R)y_R < 0$ and $q'(\theta_I) < 0$ due to the definitions of the matching functions, and that $\frac{d\theta_R}{d\theta_I} > 0$ from the definition $\theta_R = \frac{p(\theta_1)p(\theta_I(\theta_1))+\theta_1(1-q(\theta_1))}{1-p(\theta_1)}$. Thus, there is a unique θ_I for a given θ_1.

Furthermore, $V_I(\theta_I;\theta_1)$ is strictly decreasing in θ_1 following from the definition of θ_R and that $\frac{\partial V_R}{\partial \theta_R} < 0$, which implies that $\theta_I(\theta_1)$ is strictly decreasing.

Next consider the case $\theta_1 > \theta_1^1$. Then V_R is so small that $V_I(\theta_I;\theta_1)$ is negative, in which case $\theta_I(\theta_1)$ is a constant function equal to zero. Finally note, $V_I(\theta_I;\theta_1)$ is continuous in θ_1.

Then, to show existence and uniqueness of the overall equilibrium define

$$V_1(\theta_1) \equiv q(\theta_1)(1-\epsilon)[y_1 + y_2 + \max_{\theta_I, w_I \ s.\ to\ V_I=0} \{p(\theta_I)[V_R + w_I - y_2]\} - V_R - U_2] + V_R - c - K. \quad (36)$$

First, let K^l be the solution to $V_1(\theta_1^1) = q(\theta_1^1)(1-\epsilon)[y_1+y_2] + (1-q(\theta_1^1))V_R - c - K^l = 0$. Hence $K > K^l$ together with condition (35) ensures that all submarkets are active. Let K^h be the solution to $V_1(0) = (1-\epsilon)[y_1+y_2] - c - K^h = 0$, which implies the level of innovation cost leading to zero profits given the lowest possible profit of the innovating firm. Clearly, $K^h > K^l$. We assume that $K \in (K^l, K^h)$.

Note that a complication in our setup is that the matching probabilities have flat regions so that at $q = 1$ or $p = 1$, the elasticities w.r.t. θ is not defined, while they are ϵ and $1-\epsilon$ otherwise. This give rise to trivial but cumbersome restrictions. When setting K^h this is taken care of by that at $q_1 = 1$ the firm gets a share $(1-\epsilon)$ of the surplus. If the share is larger at this point, more firms would enter at K^h and we would back in the interior with $q_1 < 1$ a share $(1-\epsilon)$, and thus negative profits. $K < K^h$ thus ensures that profits are high enough to not be on the boundary $q_1 = 1$.

Now, by the properties of the matching technology we have $\lim_{\theta_1 \to \infty} V_1(\theta_1) = -c - K$, hence, by continuity, equation (36) has at least 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, first note that

$$\begin{aligned} &\frac{d}{d\theta_1}[y_1 + y_2 + \max_{\theta_I, w_I \ s.\ to\ V_I=0} \{p(\theta_I)[V_R + w_I - y_2]\} - V_R - U_2] \\ &= \frac{dV_R}{d\theta_1}(p(\theta_I) - 1) - \frac{dU_2}{d\theta_1}, \end{aligned}$$

where we have used the envelope theorem. It then follows from (36) that

$$\begin{aligned} \frac{\partial V_1(\theta_1)}{\partial \theta_1} =& q'(\theta_1)(1-\epsilon)[y_1 + y_2 + \max_{\theta_I, w_I \ s.\ to\ V_I=0} \{p(\theta_I)[V_R + w_I - y_2]\} - V_R - U_2] + \\ & q(\theta_1)(1-\epsilon)[\frac{dV_R}{d\theta_1}(p(\theta_I) - 1) - \frac{dU_2}{d\theta_1}] + \frac{dV_R}{d\theta_1} \\ =& q'(\theta_1)(1-\epsilon)[y_1 + y_2 + \max_{\theta_I, w_I \ s.\ to\ V_I=0} \{p(\theta_I)[V_R + w_I - y_2]\} - V_R - U_2] + \\ & \frac{dV_R}{d\theta_1}[1 + q(\theta_1)(1-\epsilon)(p(\theta_I) - 1)] - q(\theta_1)(1-\epsilon)\frac{dU_2}{d\theta_1}. \end{aligned}$$

Last, the facts $\frac{\partial q(\theta_1)}{\partial \theta_1} < 0$, $\frac{\partial V_R}{\partial \theta_R} < 0$, $\frac{\partial U_2}{\partial \theta_R} > 0$, and $1 + q(\theta_1)(1-\epsilon)(p(\theta_I) - 1) > 0$ imply that it suffices to show that θ_R is increasing in θ_1 to show that $\frac{\partial V_1(\theta_1)}{\partial \theta_1} < 0$. Suppose not, i.e., suppose θ_I falls so much that θ_R decreases. Then V_R increases so that $w_2 = y_2 - V_R$ must decrease. But then it follows from the zero-profit condition for the imitating firms (14) that θ_I increases, which is a contradiction. This proves the result.

10.4 Proof of Lemma 1

The first-order condition of the firm's period-2 problem of maximizing (21) reads

$$\frac{dJ_2}{dw_2} = \hat{p}_I(w_2) - 1 + \frac{d\hat{p}_I(w_2)}{dw_2}[V_R + w_2 - y_2] = 0. \tag{37}$$

Substituting in the expression for $\frac{d\hat{p}_I(w_2)}{dw_2}$ from (33) and solving for w_2 gives

$$w_2 = y_I - \frac{\hat{p}_I(w_2)(1-\epsilon)}{\hat{p}_I(w_2) - \epsilon}\big(y_I - y_2 + V_R\big),$$

where $\hat{p}_I(w_2)$ is determined by the zero profit condition of the imitating firms:

$$q(\hat{\theta}_I(w_2)) = \frac{c}{(1-\epsilon)(y_I - w_2)}. \tag{38}$$

We can combine the two equations to:

$$q(\hat{\theta}_I(w_2)) = \frac{c}{(1-\epsilon)^2(y_I - y_2 + V_R)}[1 - \frac{\epsilon}{p(\hat{\theta}_I(w_2))}].$$

The left hand side is decreasing in θ_I , starting from 1 at $\theta_I = 0$ and approaching 0 as $\theta_I \to \infty$. The right hand is non-negative only if $p(\theta_I) \geq \epsilon$ as by assumption we must have $y_I - y_2 + V_R > 0$. For θ_I above this threshold, the right-hand side is increasing until $p(\theta_I) = 1$, where it reaches the value $c/[(1-\epsilon)(y_I - y_2 + V_R)] > 0$. Thus, a unique intersection exists. Note that the intersection point moves to the left as V_R decreases so that $\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon}$ decreases. Hence, w_2 increases in θ_R. It is easy to show that for given θ_R the interior wage w_2 is smaller than in the full commitment case.

If the wage w_2 is at its lower bound U_2, it increases in θ_R since U_2 increases in θ_R. That $U_2 < y_2 - V_2$, for given θ_R, follows from the definitions of the terms together with the fact that $y_2 > \epsilon y_2 = w_R$.

Finally, we establish sufficiency at the interior solution for w_2. We need to show that

$$\frac{d^2 J_2}{(dw_2)^2} = \frac{d^2\hat{p}_I(w_2)}{(dw_2)^2}[V_R + w_2 - y_2] + 2\frac{d\hat{p}_I(w_2)}{dw_2} < 0.$$

Using the expression for $\frac{d\hat{p}_I(w_2)}{dw_2}$ from (33) and the second derivative analogous to (34):

$$\frac{d^2\hat{p}_I(w_2)}{(dw_2)^2} = -\frac{1-\epsilon}{\epsilon^2}(2\epsilon-1)\frac{\hat{p}_I(w_2)}{(y_I-w_2)^2},$$

we get

$$\begin{aligned}\frac{d^2 J_2}{(dw_2)^2} &= -\frac{1-\epsilon}{\epsilon}\frac{\hat{p}_I(w_2)}{y_I-w_2}\left[(\frac{2\epsilon-1}{\epsilon})(\frac{V_R-y_2+w_2}{y_I-w_2}-1)+2\right]<0\\ &\Leftrightarrow (2\epsilon-1)(V_R-y_2+w_2)+2\epsilon(y_I-w_2)>0\\ &\Leftrightarrow 2\epsilon(V_R-y_2+y_I)>w_2+V_R-y_2.\end{aligned}$$

As argued above, when the imitation market is active, we have $y_2 - V_R < y_I$ and the wage $w_2 < y_2 - V_R$. Thus it follows that the left hand side is larger zero, and the that left hand side is less than zero. This completes the proof.

10.5 Proof of Proposition 2

First consider the parameter bounds. Let θ_1^{1llc} denote the highest value of θ_1 such that imitating firms are indifferent between entering and not entering the market when no other firms enter to imitate. By contradiction, assume that $\theta_1^{1llc} \leq \theta_1^1$, where θ_1^1 is defined in the proof of Proposition 1. Then $V_R^{1lc} \geq V_R^1 > y_2 - y_I$. But then the wage is $w_2^{1lc} < w_2^1$ and imitating firms make strictly positive profits. Thus, we must have $\theta_1^{1lc} > \theta_1^1$. It follows that K^{llc}, given by $q(\theta^{1lc})(1-\epsilon)(y_1 + y_2) + (1 - q(\theta^{1lc}))V_R - K^{llc} - c = 0$, is smaller than K^l from the full commitment case. Also with limited commitment, the value of entering at $\theta_1 = 0$ is at least $(1-\epsilon)(y_1 + y_2) - c - K$. Hence K^h is the same in the commitment and no-commitment case, and the parameter set of the full commitment case is contained in the one for the limited commitment case.

We first establish the following lemma:

Lemma 3 *For given θ_1, define $\theta_I(\theta_1)$ as the period-2 equilibrium value of θ_I, i.e., the value that is consistent with optimal wage setting of innovating firms and imitating firms, and the zero profit condition of imitating firms. Then $\theta_I(\theta_1)$ is a decreasing, continuous function. Analogously, define $\theta_R = \theta_R(\theta_1)$ as the period-2 equilibrium value of θ_R as a function of θ_1. Then $\theta_R(\theta_1)$ is a decreasing, continuous function.*

Proof. From Lemma 1 we know that w_2^{lc} is increasing in θ_R. It follows from the zero profit condition of the imitating firm that θ_I is decreasing in θ_R. Hence we can write θ_I as a decreasing function of θ_R. Define $\hat{\theta}_I(\theta_R)$ as this function. The equilibrium value of θ_R thus writes (from 7)

$$\theta_R = \frac{p(\theta_1)p(\hat{\theta}_I(\theta_R)) + \theta_1(1-q(\theta_1))}{1-p(\theta_1)}. \tag{39}$$

The left-hand side is increasing and the right-hand side decreasing in θ_R, and hence the solution, if it exists, is unique. As the right-hand side of the equation is bounded upwards, the left-hand side is strictly greater than the right-hand side for θ_R sufficiently high. The right-hand side of the equation reaches its minimum for $\theta_I = 0$, where the right-hand side is equal to $\frac{\theta_1(1-q(\theta_1))}{1-p(\theta_1)}$, which ensures existence. Note that even though $\theta_I(\theta_1)$ in principle may be zero, our parameter assumptions ensure that this is not the case for values of θ_1 close to the equilibrium value.

We next show that $\theta_R(\theta_1)$ is increasing in θ_1. The proof is by contradiction. Suppose $\theta_R(\theta_1)$ is nonincreasing in θ_R. It then follows that θ_I in equilibrium is nondecreasing since $\hat{\theta}_I(\theta_R)$ is a decreasing function of θ_R. Last it follows from (7) that θ_R is strictly increasing in θ_1, a contradiction.

As indicated in footnote 21 the function $w_2^{lc}(\theta_R)$ is discontinuous at one point, at which point w_2^{nc} jumps from U_2 to the interior solution $\hat{w}_2$. Hence, the contraction (39) may not have a solution, in which case the period-2 model has no equilibrium in pure strategies. In this case we assume that firms randomize between setting $w_2 = U_2$ and $w_2 = \hat{w}_2$. To be more specific, suppose the point of discontinuity is at $\bar{\theta}_R$, and suppose the right-hand side of (39) is $\theta^h > \bar{\theta}_R$ when all firms set $w_2 = U_2$, and $\theta^l < \bar{\theta}_R$ when all firms set $w_2 = \hat{w}_2$. Then there exists a probability π such that if a fraction π sets $w_2 = U_2$ and a fraction $1 - \pi$ sets $w_2 = \hat{w}_2$, then $\bar{\theta}_R$ emerges. We then set $\theta_R(\theta_1) = \bar{\theta}_R$. On the interval where the firms randomize, and $\theta_R(\theta_1)$ is constant. Hence $V_2(\theta_1)$ is constant at this interval, while M_2 is increasing in θ_1 at this interval.

■

Proof of Existence: First note that in the same manner as in the proof of Proposition 1 we can pick parameter values such that the imitation market is open. In the following we denote $V_1^{**}(\theta_1)$ the equilibrium value of V_1 given θ_1. Using the functions from Lemma 3, the value of the innovating firm in the first period can then be written:

$$\begin{aligned}
V_1^{**}(\theta_1) &= q(\theta_1)(1-\epsilon)[y_1 + y_2 + p(\theta_I)[V_R + w_I - y_2] - V_R - U_2] + V_R - (K + c) \\
&= q(\theta_1)(1-\epsilon)[y_1 + y_2 + p(\theta_I(\theta_1))[q(\theta_R(\theta_1))(1-\epsilon)y_R + \epsilon y_I - y_2 + (1-\epsilon)(w_2^{nc}(\theta_R(\theta_1))y)] \\
&\quad - q(\theta_R(\theta_1))(1-\epsilon)y_R - p(\theta_R(\theta_1))\epsilon y_R] + q(\theta_R(\theta_1))(1-\epsilon)y_R - (K + c),
\end{aligned}$$

when the lower bound on w_I does not bind. We have to show that a solution to $V_1^{**}(\theta_1) = 0$ exists. First, note even though it may be a discontinuity where $w_2^{nc}(\theta_R)$ jumps, $V_1^{**}(\theta_1)$ is continuous since θ_R is constant (see the proof of Lemma 3 for details). Also, by assumption we have that $V_1^{**}(0) \geq (1-\epsilon)(y_1 + y_2) - K - c > 0$. 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. Existence when the lower bound on w_I binds follows from a similar argument and is omitted.

Proof of Uniqueness: We have to show that V_1 defined by (19) is decreasing in θ_1. Taking derivatives give

$$\frac{dV_1}{d\theta_1} = q'(\theta_1)(1-\epsilon)[M_1 - V_R - U_2] + q(\theta_1)(1-\epsilon)\frac{d}{d\theta_1}[M_1 - U_2] + (1 - q(\theta_1)(1-\epsilon))\frac{dV_R}{d\theta_1}.$$

The first term is strictly negative. From Lemma 3 we know that θ_R is decreasing in θ_1. Hence the third term is negative. We are left with the second term. More specifically, we have to show that

$M_1 - U_2$ is decreasing in θ_1, or equivalently that $M_2 - U_2$ is decreasing in θ_1. Recall that M_2 can be written as

$$M_2 = p_I(w_I + V_R - y_2) + y_2 \tag{40}$$

Suppose first that the lower bound on w_I binds, i.e., that $V_R = y_2 - w_I$. It follows that $M_2 = y_2$. Since U_2 is increasing in θ_R and hence in θ_1, it follows that $M_2 - U_2$ is decreasing, and hence that $dV/d\theta_1 < 0$.

Suppose then that the lower bound w_2 binds, i.e., that $w_2 = U_2$, while the lower bound on w_I does not. Then $w_I = \epsilon y_I + (1-\epsilon)U_2$ and we can write

$$\begin{aligned} M_2 - U_2 &= p(\theta_I)(\epsilon y_I + (1-\epsilon)U_2 + V_R - y_2) + y_2 - U_2 \\ &= (p(\theta_I)(1-\epsilon) - 1)U_2 + p(\theta_I)(\epsilon y_I + V_R - y_2) + y_2 \end{aligned}$$

Since U_2 is increasing in θ_1, and θ_I is decreasing in θ_1, it is straightforward to show that this expression is decreasing in θ_1.

Suppose then that $w_2 = \hat{w}_2 > U_2$. By taking derivative of M_2 in (40) it follows that

$$\frac{dM_2}{d\theta_1} = p'(\theta_I)\theta_I'(\theta_1)[w_I + V_R - y_2] + p(\theta_I)[\frac{dw_I}{d\theta_1} + \frac{dV_R}{d\theta_1}].$$

Since $w_I + V_R - y_2 > 0$, $\theta_I'(\theta_1)$ is negative and $p'(\theta_I)$ is positive, the first term is negative. Hence a sufficient (but not necessary) condition for M_2 to be decreasing in θ_1 is that $\frac{dw_I}{d\theta_1} + \frac{dV_R}{d\theta_1} < 0$. Since $w_I = \epsilon y_I + (1-\epsilon)w_2$, a sufficient (but not necessary) condition for this to hold is that $\frac{dw_2}{d\theta_1} + \frac{dV_R}{d\theta_1} < 0$. We will now derive sufficient conditions for this to hold. To that end, recall from (37) that the first-order condition for w_2 reads

$$-(1 - \hat{p}_I(w_2)) + \hat{p}_I'(w_2)[w_2 + V_R - y_2] = 0.$$

Taking derivative with respect to θ_1 gives

$$\hat{p}_I'(w_2)\frac{dw_2}{d\theta_1} + \hat{p}_I''(w_2)\frac{dw_2}{d\theta_1}[w_2 + V_R - y_2] + \hat{p}_I'(w_2)\frac{d}{d\theta_1}[w_2 + V_R - y_2] = 0.$$

Hence

$$\frac{d}{d\theta_1}[w_2 + V_R] = \frac{-\hat{p}_I'(w_2) + \hat{p}_I''(w_2)[y_2 - w_2 - V_R]}{\hat{p}_I'(w_2)}\frac{dw_2}{d\theta_1}.$$

We know that $\frac{dw_2}{d\theta_1} > 0$ and that $\hat{p}_I'(w_2) < 0$. Hence the first term in the nominator is positive, while the denominator is negative. Recall that $y_2 - w_2 - V_R > 0$. Hence a sufficient condition for the right hand side to be negative is that $\hat{p}_2''(w_2) > 0$. Recall from (38) that

$$q(\theta_I) = k_1(y_I - w_2)^{-1},$$

where k_1 is a constant. Using the definition of $q(\theta_I)$, it follows that $\theta_I = k_2(y_I - w_2)^{\frac{1}{\epsilon}}$ and we can write

$$\hat{p}_2(w_2) = k_3(y_I - w_2)^{\frac{1-\epsilon}{\epsilon}},$$

where k_2 and k_3 are uninteresting constants. Then, taking derivative twice gives

$$\hat{p}_2''(w_2) = \frac{1-\epsilon}{\epsilon}\frac{1-2\epsilon}{\epsilon}k_3(y_I - w_2)^{\frac{1-3\epsilon}{\epsilon}}.$$

It follows that $\hat{p}''(w_2) > 0$ if and only if $\epsilon < 1/2$. The result thus follows.

10.6 Proof of Proposition 3

We first prove the following Lemma.

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 arguments 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, $w_2^*(\theta_1) = y_2 - V_R^*(\theta_1) \geq y_2 - V_R^{**}(\theta_1) > w_2^{**}(\theta_1)$, and hence, by the zero-profit condition of the imitating firms, $\theta_I^*(\theta_1) < \theta_I^{**}(\theta_1)$, a contradiction. Further, given $\theta_I^*(\theta_1) < \theta_I^{**}(\theta_1)$, zero-profit condition implies $w_2^{**}(\theta_1) < w_2^*(\theta_1)$.

When w_I is bounded, i.e. $w_I = y_2 - V_R$, the zero-profit condition is given by

$$q(\theta_I) = \frac{c}{(y_I - y_2 + V_R)}. \tag{41}$$

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 by (41) $\theta_I^*(\theta_1) < \theta_I^{**}(\theta_1)$, which is a contradiction. ■

To prove Proposition 3, insert for U_2 in equation (36) for the full commitment case to get

$$V_1^*(\theta_1) = q(\theta_1)(1-\epsilon)[y_1 + y_2 + \max_{\theta_I, w_I \ s.\ to\ V_I=0}\{p(\theta_I)[V_R + w_I - y_2]\} - p(\theta_R)\epsilon y_R] - K - c.$$

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:

$$\max_{\theta_I, w_I \ s.\ to\ V_I=0}\{p(\theta_I)[V_R + w_I - y_2]\} \geq p(\theta_I^{**}(\theta_1))[V_R + w_I^{**}(\theta_1) - y_2].$$

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 $\theta_1^* > \theta_1^{**}$.

When w_I is bounded we can write the profits of the innovating firms in equilibrium as:

$$\begin{aligned} V_1^{**}(\theta_1) &= q(\theta_1)(1-\epsilon)[y_1 + y_2 + p(\theta_I)[V_R + \check{w}_I - y_2] - p(\theta_R)\epsilon y_R] - K - c \\ &= q(\theta_1)(1-\epsilon)[y_1 + y_2 - p(\theta_R)\epsilon y_2] - K - c. \end{aligned}$$

Following a similar argument as above, showing existence is straightforward. To show Proposition 3 when $w_I = \check{w}_I$, first note that in the full commitment case $\max_{\theta_I, w_I \ |V_I=0}\{p(\theta_I)[V_R + w_I - y_2]\} > 0$. 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 argument as above, we can conclude that $\theta_1^* > \theta_1^{**}$.

10.7 Proof of Proposition 4

First we compare the first-order condition $\frac{\partial F}{\partial \theta_1} = 0$ to the zero profit condition $V_1 = 0$. Taking the derivative of the welfare function in (23) with respect to θ_1 and using $\theta_R = \frac{p(\theta_I)p(\theta_1)+\theta_1(1-q(\theta_1))}{1-p(\theta_1)}$ and the fact that $p'(\theta_i) = (1-\epsilon)q(\theta_i)$ we get

$$\frac{\partial F}{\partial \theta_1} =$$
$$q(\theta_1)(1-\epsilon)[y_1 + y_2 + p_I(y_I - y_2) - c\theta_I] + \frac{d}{d\theta_1}[(1-p(\theta_1))p(\theta_R)y_R] - (c+K) - c_R[(1-\epsilon)q(\theta_1)(p(\theta_I)-1)+1]$$

Using the free entry condition for the imitating firms we can replace c and rewrite the first summand as:

$$q(\theta_1)(1-\epsilon)[y_1 + y_2 + p(\theta_I)(w_I - y_2)].$$

Next, we turn to the second summand:

$$\begin{aligned} &\frac{d}{d\theta_1}[(1-p(\theta_1))p(\frac{p(\theta_I)p(\theta_1)+\theta_1(1-q(\theta_1))}{1-p(\theta_1)})y_R] \\ &= -(1-\epsilon)q(\theta_1)p(\theta_R)y_R + (1-p(\theta_1))p'(\theta_R)y_R[\frac{(1-\epsilon)q(\theta_1)(p(\theta_I)p(\theta_1)+\theta_1(1-q(\theta_1)))}{(1-p(\theta_1))^2} \\ &+ \frac{(1-\epsilon)q(\theta_1)(p(\theta_I)-1)+1}{1-p(\theta_1)}] \\ &= (1-\epsilon)q(\theta_1)[-\epsilon p(\theta_R) - (1-p(\theta_I))(1-\epsilon)q(\theta_R)y_R)] + (1-\epsilon)q(\theta_R)y_R) \\ &= (1-\epsilon)q(\theta_1)[-U_2 - (1-p(\theta_I))(V_R + c_R)] + V_R + c_R. \end{aligned}$$

Combining we get:

$$\frac{\partial F}{\partial \theta_1} = (1-\epsilon)q(\theta_1)[y_1 + y_2 + p(\theta_I)(w_I - y_2) - U_2 - (1-p(\theta_I))(V_R)] + V_R - (c+K) = 0.$$

Now we will compare this first-order condition to the zero-profit condition for innovating firms, which can be written as

$$V_1 = q(\theta_1)(1-\epsilon)[M_1 - U_2 - V_R] + V_R - (c+K) = 0.$$

Substituting in the definition of M_1 as given in (16) we get the desired result.

Next, the first-order condition with respect to θ_I can be written as

$$\frac{\partial F}{\partial \theta_I} = p(\theta_1)[p'(\theta_I)(y_I - y_2) - c] + (1-p(\theta_1))p'(\theta_R)\frac{d\theta_R}{d\theta_I}y_R - p(\theta_1)p'(\theta_I)c_R = 0.$$

We have

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

Using the definition of $V_R = (1-\epsilon)q(\theta_R)y_R - c_R$ and dividing by $p(\theta_1) > 0$ we get

$$(1-\epsilon)q(\theta_I)(y_I - y_2 + V_R) - c = 0,$$

which is identical to the zero profit condition of imitating firms given in (20).

10.8 Proof of Lemma 2

For the first part, i.e. $\frac{\partial F(\theta_1^{**},\theta_I^{**})}{\partial \theta_1} = 0$ we can just refer to the first part of the efficiency result (shown in appendix 10.7). We can apply that proof as the equality of the first order condition and the zero profit condition holds for any given level of θ_I. To see this note that in effect, when the firm chooses w_1 in period 1, it takes M_1 as given and maximizes $V_1 = q(\theta_1)(M_1 - W_1) + (1-q(\theta_1))V_R - c - K$ subject to (3), which gives the first-order condition $W_1 = \epsilon(M_1 - V_R) + (1-\epsilon)U_2$. Using this to substitute out W_1 from V_1 gives the free entry condition

$$V_1 = q(\theta_1)(1-\epsilon)[M_1 - V_R - U_R] + V_R - (c+K) = 0,$$

which we have shown in appendix 10.7 is the same as the first-order condition for efficiency.

Next, we establish the second condition, $\frac{\partial F(\theta_1^{**},\theta_I^{**})}{\partial \theta_I} < 0$. The derivative of the welfare function with respect to θ_I is

$$\frac{\partial F(\theta_1^{**},\theta_I^{**})}{\partial \theta_I} = p(\theta_1)[p'(\theta_I)(y_I - y_2) - c] + (1-p(\theta_1))p'(\theta_R)\frac{d\theta_R}{d\theta_I}y_R - p(\theta_1)p'(\theta_I)c_R.$$

We have

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

Substituting out $c = q(\theta_I)(y_I - w_I)$ and using the result $w_I = \epsilon y_I + (1-\epsilon)w_2$ we get:

$$\frac{\partial F(\theta_1^{**}, \theta_I^{**})}{\partial \theta_I} = p(\theta_1)[(1-\epsilon)^2 q(\theta_I)q(\theta_R)y_R - (1-\epsilon)q(\theta_I)(y_2 - w_2) - (1-\epsilon)q(\theta_I)c_R].$$

Last, use the definition of $V_R = (1-\epsilon)q(\theta_R)y_R - c_R$ to get:

$$\frac{\partial F(\theta_1^{**}, \theta_I^{**})}{\partial \theta_I} = (1-\epsilon)p(\theta_1)q(\theta_I)[V_R - y_2 + w_2] < 0,$$

since $w_2 > y_2 - V_R$ in the limited commitment equilibrium. 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 or w_I are bound follow exactly the same line of argument and are therefore omitted).

10.9 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 2. Then, the inequality follows from the fact that higher θ_1 implies lower θ_I as stated in Lemma 3.

10.10 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, $q_I(\theta_I) = c/(y_I - w_I)$, 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. Then by the zero-profit condition of the imitating firms, w_2 must decrease. Since θ_I increases, θ_1 has to decrease sufficiently to lower θ_R in order for w_2 (as given in (22)) to go down. The zero-profit condition of the innovators can be written

$$q(\theta_1)(1-\epsilon)(M_1 - U_2) + V_R[1 - q(\theta_1)(1-\epsilon)] - K + c = 0, \tag{42}$$

where

$$\begin{aligned} M_1 =& y_1 + y_2 + p(\theta_I)[V_R + \epsilon y_I + (1-\epsilon)w_2 - y_2] \\ =& y_1 + y_2 + p(\theta_I)[\epsilon y_I - y_2 + (1-\epsilon)\frac{(1-\epsilon)p(\theta_I)y_2 - \epsilon(1-p(\theta_I))y_I}{p(\theta_I) - \epsilon} \\ &+ \left(1 - (1-\epsilon)\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I) - \epsilon}\right)V_R], \end{aligned}$$

when the w_I is not bound, i.e. $w_I + V_R > y_2$. First, consider the effect of $d\theta_I$. The sign of the inner part of $\frac{dM_1}{d\theta_I}$ is then given by

$$\frac{\partial\left((1-\epsilon)\frac{(1-\epsilon)p(\theta_I)y_2-\epsilon(1-p(\theta_I))y_I-(1-\epsilon)\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon}V_R}{p(\theta_I)-\epsilon}\right)}{\partial p(\theta_I)}\frac{dp(\theta_I)}{d\theta_I}.$$

Since θ_I increases, this is positive as $(1-\epsilon)\frac{\partial w_2}{\partial p(\theta_I)} = (1-\epsilon)^2\frac{\epsilon(y_I+V_R-y_2)}{(p(\theta_I)-\epsilon)^2} > 0$ (since $y_I + V_R > y_2$) and $\frac{dp(\theta_I)}{d\theta_I} > 0$. Next, the effect on profits given by (42) of a change in θ_R reads

$$q(\theta_1)(1-\epsilon)(\frac{dM_1}{d\theta_R} - \frac{dU_2}{d\theta_R}) + \frac{\partial V_R}{\partial\theta_R}[1-q(\theta_1)(1-\epsilon)],$$

where $\frac{dM_1}{d\theta_R} = q(\theta_1)(1-\epsilon)\left(1-(1-\epsilon)\frac{p(\theta_I)(1-\epsilon)}{p(\theta)-\epsilon}\right)\frac{\partial V_R}{\partial\theta_R}$. Notice that U_2 decreases as θ_R decrease. Thus, since V_R is decreasing in θ_R, to show that profits increase as θ_R decreases it is sufficient to show that

$$q(\theta_1)(1-\epsilon)\left(1-(1-\epsilon)\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon}\right) + [1-q(\theta_1)(1-\epsilon)] > 0,$$

which follows from that $1-(1-\epsilon)\frac{p(\theta_I)(1-\epsilon)}{p(\theta_I)-\epsilon} > 0$ when $w_I + V_R > y_2$. Therefore profits increase when θ_I increases and θ_R decreases.

Finally, since $M_1 - V_R - U_2 > 0$, 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 2.

10.11 Proof of Derivative of $\hat{F}$ w.r.t. ρ

We will go through each term of 30 in turn. First, from Lemma 2 we know that the first term is zero. Second,

$$\begin{aligned}
\frac{\partial\hat{F}}{\partial\theta_I} =& p(\theta_1)\frac{d}{d\theta_I}\left\{(1-\rho)p(\theta_I)[w_I + q(\frac{p(\theta_1)(1-\rho)p_I(\theta_I)+\theta_1(1-q(\theta_1))}{1-p(\theta_1)})y_R - y_2 - c_R]\right\} \\
&+ \frac{d}{d\theta_I}\left\{\theta_1(1-q(\theta_1))q(\frac{p(\theta_1)(1-\rho)p_I(\theta_I)+\theta_1(1-q(\theta_1))}{1-p(\theta_1)})\right\} \\
=&(1-\rho)p(\theta_1)[w_I + q(\theta_R)y_R(1-\frac{\epsilon p_I(\theta_I)}{\theta_R}\frac{p(\theta_1)(1-\rho)}{1-p(\theta_1)}) - y_2 - c_R]\frac{dp(\theta_I)}{d\theta_I} \\
&- \theta_1(1-q(\theta_1))\frac{\epsilon q(\theta_R)y_R}{\theta_R}\frac{p(\theta_1)(1-\rho)}{1-p(\theta_1)}\frac{dp(\theta_I)}{d\theta_I} \\
=&(1-\rho)p(\theta_1)[w_I + V_R - y_2]\frac{dp(\theta_I)}{d\theta_I} > 0,
\end{aligned}$$

where we have used the fact that $V_R = (1-\epsilon)q(\theta_R)y_R - c_R$. The inequality follows from the wage bound 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 (30) is strictly negative.

From (29) it follows that $\partial \hat{F}/\partial w_I = p(\theta_1)(1-\rho)p(\theta_I) > 0$. In the following Lemma we show that $dw_I/d\rho < 0$.

Lemma 5 *It holds that* $\frac{dw_I}{d\rho} < 0$.

Proof. 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 it follows from the equilibrium value of w_2 that w_2 (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:

$$q(\theta_1)(1-\epsilon)(M_1 - V_R - U_2) + V_R = K + c, \tag{43}$$

where

$$M_1 = y_1 + y_2 + (1-\rho)p(\theta_I)(w_I + V_R - y_2).$$

Suppose θ_1 increases so much that θ_R increases enough so that w_2 stays constant. Then it follows that M_1 falls if $w_I + V_R > y$ as V_R is decreasing in θ_R. Furthermore, U_2 increases and, since $M_1 - V_R - U_2 > 0$, the left hand side of (43) then decreases. Given that the equilibrium is locally stable, it follows that θ_1 cannot increase that much, hence the result follows. The case of $w_I = y_2 - V_R$ can be established in a similar way. ■

Hence the third term in (30) is also negative. Finally, we have that

$$\begin{aligned}
\frac{\partial \hat{F}}{\partial \rho} =& p(\theta_1)\frac{d}{d\rho}\left\{(1-\rho)p(\theta_I)[w_I + q(\frac{p(\theta_1)(1-\rho)p_I(\theta_I) + \theta_1(1-q(\theta_1))}{1-p(\theta_1)})y_R - y_2 - c_R]\right\} \\
&+ \frac{d}{d\rho}\left\{\theta_1(1-q(\theta_1))q(\frac{p(\theta_1)(1-\rho)p_I(\theta_I) + \theta_1(1-q(\theta_1))}{1-p(\theta_1)})y_R\right\} \\
=& -p(\theta_1)p(\theta_I)[w_I + V_R - y_2] < 0,
\end{aligned}$$

where the steps are similar to the steps for $\partial \hat{F}/\partial \theta_I$.

10.12 Proof of $dM_1/d\rho < 0$

Using the definitions of J_2 in (21) and W_2 in (2) we can write

$$M_1 = J_2 + W_2 + y_1.$$

From the firm's perspective, θ_R is given. From the envelope theorem it follows that have $\frac{dJ_2}{d\rho} = -p(\theta_I)[V_R + w_2 - y_2] > 0$ since $w_2 < y_2 - V_R$. Furthermore, the envelope theorem also implies that (since the imitation market maximizes the income of the searching workers) $\frac{dW_2}{d\rho} = \frac{dw_2}{d\rho}(1 + (1 - \rho)p(\theta_I)) - p(\theta_I)(w_I - w_2)$. Combining gives

$$\frac{dM_1}{d\rho} = -p(\theta_I)[V_R + w_I - y_2] + \frac{dw_2}{d\rho}(1 + (1 - \rho)p(\theta_I)),$$

where the first part is negative due to the bound $w_I \geq y_2 - V_R$. What is left to show is that $\frac{dw_2}{d\rho} \leq 0$.

Let $w_2(\rho)$ be the innovating firm's optimal period-2 wage as a function of ρ. If the wage is bound by U_2, the result is immediate θ_R is given from the firm's perspective. For the interior period-2 wage, the first-order condition is the solution to the equation (derived analogously as in appendix 10.4):

$$w_2(\rho) = y_I - \frac{(1 - \rho)\hat{p}_I(w_2(\rho))(1 - \epsilon)}{(1 - \rho)\hat{p}_I(w_2(\rho)) - \epsilon}[y_I - y_2 + V_R].$$

Taking derivative with respect to ρ gives

$$\frac{dw_2(\rho)}{d\rho} = -\frac{\epsilon(1 - \epsilon)\hat{p}_I(w_2(\rho))[y_I - y_2 + V_R]}{(1 - \rho)(\hat{p}_I(w_2(\rho)) - \epsilon)^2} + \frac{(1 - \rho)(1 - \epsilon)\epsilon[y_I - y_2 + V_R]}{((1 - \rho)\hat{p}_I(w_2(\rho)) - \epsilon)^2}\frac{d\hat{p}_I(w_2(\rho))}{d\rho}.$$

To show the result, suppose the opposite is true, i.e. $\frac{dw_2}{d\rho} \geq 0$. Note that $y_I - y_2 + V_R > 0$ so the first term is negative. Then$\frac{dw_2}{d\rho}$ can only be positive if $\frac{\hat{p}_I(w_2(\rho))}{d\rho} > 0$. However, from the imitating firm's zero-profit condition, $(1 - \rho)q(\theta_I)(1 - \epsilon)(y_I - w_2) = c$, it follows that a higher ρ in tandem with a higher wage w_2 certainly means a lower θ_I and hence a lower p. a contradiction.