8 Appendix

8.1 Marginal Monitoring Costs

Regarding the costs of monitoring, we assume in our main framework that these costs are borne by firms in the form of higher overhead costs. Lazear (1986) assumes that monitoring creates higher marginal costs that are passed on to workers. This is an alternative way of modelling these costs that can also be integrated in our framework and leads to qualitatively identical results.

Suppose workers in adopting firms can only spent a fraction $\chi \in (0,1)$ on their production tasks and $1-\chi$ is spent on reporting their progress (monitoring). This implies that adopting firms do not have to pay a higher fixed costs in order to monitor, but a variable iceberg cost per worker. Marginal unit labor requirements become $\chi^{-1} > 1$, and total variable labor requirements in adopting firms are $\chi^{-1} x_a$. Fixed costs are the same for both types of firms: $f_a = f_n = f$, so that $\Phi = 1$.

Following Lazear (1986), monitoring firms pass on the costs of monitoring to their workers, paying a discount on wages by a factor of χ. Hence, wages paid by adopting firms are now $w_a(z) = \chi c_a e(\bar{\varepsilon}, z)$. As a consequence, their profits are given by $\pi_a = A(\chi c_a/\chi)^{1-\sigma} - f = A c_a^{1-\sigma} - f$. Non-monitoring firms continue to pay wages $w_n = c_n \bar{z}_n(\tilde{z})$ and earn profits $\pi_n = A c_n^{1-\sigma} - f$. Free entry leads to zero profits of both types of firms and equalizes effective wages across firms: $A c_n^{1-\sigma} = f = A c_a^{1-\sigma}$, or $c_a = c_n$.

Sorting now requires that $w_n = c_n \bar{z}_n(\tilde{z}) = w_a(\tilde{z}) = \chi c_a e(\bar{\varepsilon}, \tilde{z})$, or, given $c_a = c_n$,

$$\frac{\bar{z}_n(\tilde{z})}{e(\bar{\varepsilon}, \tilde{z})} = \chi < 1. \tag{46}$$

Again, $\bar{z}_n(\tilde{z})/e(\bar{\varepsilon}, \tilde{z}) \in (0,1)$ implies incomplete adoption where $\tilde{z} \in (\underline{z}, \infty)$.

In the social planners problem, the second constraint now becomes $\chi^{-1} n_a^* x_a^* \leq L \int_{\tilde{z}^*}^{\infty} e(\bar{\varepsilon}, z)\, dG(z)$ and the solution is $\tilde{z}^*/e(\bar{\varepsilon}, \tilde{z}^*) = \chi$. Hence, $\tilde{z} < \tilde{z}^*$, and the market solution is again inefficient.

In order to identify this distortion we can rewrite (46) as

$$\ln \bar{z}_n(\tilde{z}) - \ln \tilde{z} = [\ln e(\bar{\varepsilon}, \tilde{z}) - \ln \tilde{z}] + \ln \chi \tag{47}$$

In this case the productivity effect of monitoring depends on two counteracting effects: A positive effect of incentivizing workers to provide effort (the first term in square brackets), and a negative effect of the costs of monitoring (the second term, $\ln \chi < 0$). Because of the supermodularity of skills and effort, the positive effect dominates for highly skilled workers,

whereas the negative effect dominates for workers in the lower distribution of Z_a. With free entry, revenues are equalized across firm types, and the two productivity effects cancel out in the aggregate. Using insider econometrics, our inefficiency would then show up as a negative productivity effect *at the margin*.

8.2 Disutility of Effort and Worker Sorting

In the main text we assume that the disutility of effort is negligible. In this appendix we discuss the implications of higher values of γ for the sorting equilibrium.

A firm that has adopted the monitoring technology can observe the level of output of each worker and so can back out the effective labor supplied by each worker. If c_a is the price of an effective unit of labor facing an adopting firm, then the utility of a worker of skill z working at an adopting firm a is given by

$$U_a(z) = \max\left\{\frac{c_a}{P}e(\bar{\varepsilon}, z) - \gamma\bar{\varepsilon}, \frac{c_a}{P}z - \gamma\right\}. \tag{48}$$

Given the supermodularity of e, a worker who has chosen a job at an adopting firm will choose to work hard (provide effort $\bar{\varepsilon} > 1$) if $z \geq \hat{z}$, where $\hat{z}$ is the solution to

$$\left[e(\bar{\varepsilon}, \hat{z}) - \hat{z}\right]\frac{c_a}{P} = \gamma(\bar{\varepsilon} - 1). \tag{49}$$

At $\hat{z}$, the real reward of higher effort is just equal to the extra disutility of effort.

Now consider the choice of whether to apply for a job at a non-adopting firm or an adopting firm. If $c_n\bar{z}(\tilde{z})$ is the going wage at a non-adopting firm, where $\tilde{z}$ is the marginal worker, who would otherwise work at an adopting firm, the utility of a worker choosing between a non-adopting firm and an adopting firm is

$$U(z) = \max\left\{U_a(z), \frac{c_n\bar{z}_n(\tilde{z})}{P} - \gamma\right\}. \tag{50}$$

The cutoff skill between adopters and non-adopters, $\tilde{z}$, depends on whether the worker would choose to exert high effort ($\varepsilon = \bar{\varepsilon}$) at an adopting firm or low effort ($\varepsilon = 1$), i.e. whether $\hat{z} \gtreqless \tilde{z}$. We begin our analysis by assuming that $\hat{z} > \tilde{z}$ so that the marginal worker would provide low effort. In this case, there are no productivity differences for individual workers between the two types of firms, and the cutoff $\tilde{z}$ is the solution to

$$c_n\bar{z}_n(\tilde{z}) = c_a\tilde{z}. \tag{51}$$

Note that γ does not appear in this expression because the worker chooses to shirk

in either case and, therefore, has the same disutility of effort in adopting and non-adopting firms. If $\hat{z} > \tilde{z}$, sorting leads to three groups of workers: The most able workers are employed at adopting firms and work hard ($z > \hat{z}$), the least capable workers work at non-adopting firms and shirk ($z < \tilde{z}$), and the moderately skilled workers work at adopting firms and choose to shirk $z \in (\tilde{z}, \hat{z})$.

Let us now explore the role of γ. Using free entry (21), Φ from (22) and $\xi = c_n/P$ from (36) we can rewrite (49) to read

$$\hat{z} = e(\bar{\varepsilon}, \hat{z}) - \frac{\gamma}{\xi\Phi}(\bar{\varepsilon} - 1) \tag{52}$$

and (51) to read

$$\frac{\bar{z}_n(\tilde{z})}{\tilde{z}} = \Phi. \tag{53}$$

As shown in Proposition 1 in the paper, equation (53) has at least one stable equilibrium. We make the necessary assumptions over G that imply that this equilibrium is unique. It is evident from the assumptions on $e(\varepsilon, z)$ in section 2.1.2 that for $\gamma > 0$ there is an interior $\hat{z}$ that is unique.

Proposition 10 *There exists a γ^* such that for $\gamma > \gamma^*$, $\hat{z} > \tilde{z}$, and for $\gamma < \gamma^*$, $\hat{z} < \tilde{z}$.*

Proof. The derivative of (52) yields

$$\frac{d\hat{z}}{d\gamma} = \frac{(\bar{\varepsilon} - 1)}{\xi\Phi\left[e_{\hat{z}}(\bar{\varepsilon}, \hat{z}) - 1\right]} > 0, \tag{54}$$

where $e_{\hat{z}}(\bar{\varepsilon}, \hat{z}) \equiv \partial e(\bar{\varepsilon}, \hat{z})/\partial\hat{z} > 1$ (by the properties of $e(\cdot)$). We also have $\lim_{\gamma\to 0}\hat{z} = \underline{z}$ and $\lim_{\gamma\to\infty}\hat{z} = \infty$. As $\hat{z}$ is strictly increasing in γ over its entire range and $\tilde{z}$ is constant in γ there must exist a γ^* such that for $\gamma > \gamma^*$, $\hat{z} > \tilde{z}$, and for $\gamma < \gamma^*$, $\hat{z} < \tilde{z}$. ■

For $\gamma < \gamma^*$ ($\hat{z} < \tilde{z}$), the cutoff conditions merge into a single expression given by

$$\frac{\bar{z}_n(\tilde{z})}{e(\bar{\varepsilon}, \tilde{z})} + \frac{\gamma(\bar{\varepsilon} - 1)}{\xi e(\bar{\varepsilon}, \tilde{z})} = \Phi \tag{55}$$

The marginal worker at an adopting firm must be compensated for the disutility of higher effort provision in these firms. But since $\hat{z} < \tilde{z}$, all workers in adopting firms with $z > \tilde{z} > \hat{z}$ work hard and provide a higher effort than in non-adopting firms.

Following our proof of existence in proposition 1, we can show that

$$\lim_{\tilde{z}\to\underline{z}}\left\{\frac{\bar{z}_n(\tilde{z})}{e(\bar{\varepsilon}, \tilde{z})} + \frac{\gamma(\bar{\varepsilon} - 1)}{\xi e(\bar{\varepsilon}, \tilde{z})}\right\} = 1 + \frac{\gamma}{\xi\underline{z}}(\bar{\varepsilon} - 1) > 1$$

and

$$\lim_{\tilde{z}\to\infty}\left\{\frac{\bar{z}_n(\tilde{z})}{e(\bar{\varepsilon},\tilde{z})}+\frac{\gamma(\bar{\varepsilon}-1)}{\xi e(\bar{\varepsilon},\tilde{z})}\right\}=0$$

So existence of a sorting equilibrium is again satisfied. Further, the condition given in proposition 1 is also sufficient for this equilibrium to be stable because the additional term in the expression (55) is monotonically decreasing in the cutoff. As γ approaches zero, the second term in (55) disappears and we return to the equilibrium described in proposition 1 and equation (23):

Proposition 11 *The sorting equilibrium described in proposition 1 and equation (23) describes the limiting case where γ approaches zero.*

For the case of $\gamma > \gamma^*$, where $\hat{z} > \tilde{z}$, we need to amend the labor market clearing condition for workers at adopting firms as there are now two, rather than one, cutoff conditions. The new labor market clearing condition for the monitored labor pool is written

$$L[1-G(\tilde{z})]\,\bar{e}_a(\bar{\varepsilon},\tilde{z},\hat{z})=n_a x_a, \tag{56}$$

where $\bar{e}_a(\bar{\varepsilon},\tilde{z},\hat{z})\equiv\left(\int_{\tilde{z}}^{\hat{z}} z dG(z)+\int_{\hat{z}}^{\infty} e(\bar{\varepsilon},z)\,dG(z)\right)/[1-G(\tilde{z})]$ is the new average productivity of workers in adopting firms with two cutoffs $\tilde{z}$ and $\hat{z}$.

The new condition that differs from those in the text are given by (52), (53), and (56). From the arguments above, the impact of a reduction in γ affects the equilibrium entirely through its role on the cutoff in equation (52) and this in turn affects the allocation of resources through equation (56). Combining this with the existing labor market clearing conditions (11) and (19) as well as with (8), (21) and (22) we obtain the following expression

$$\frac{n_a}{n_n}=\Phi^{\sigma}\frac{[1-G(\tilde{z})]}{G(\tilde{z})}\frac{\bar{e}_a(\bar{\varepsilon},\tilde{z},\hat{z})}{\bar{z}_n(\tilde{z})} \tag{57}$$

Equations (52), (55) and (57) pin down the number of adopting and non-adopting firms. The following proposition is immediate:

Proposition 12 *The ratio of adopting firms to non-adopting firms is decreasing in γ.*

Proof. Since $\tilde{z}$ is determined by (55) and independent of γ, we obtain

$$\frac{d(n_a/n_n)}{d\gamma}=\Phi^{\sigma}\frac{\hat{z}-e(\bar{\varepsilon},\hat{z})}{G(\tilde{z})\,\bar{z}_n(\tilde{z})}\frac{d\hat{z}}{d\gamma}<0, \tag{58}$$

where $d\hat{z}/d\gamma>0$ by (54). ■

This is intuitive. The amount of effective labor available to adopting firms expands even though the identity of the workers has not changed as more workers are willing to work hard when the disutility of effort declines. This causes more managers to be drawn into adopting firms.

Put differently, when γ rises, fewer and fewer workers find it profitable to work harder. This reduces the average productivity of workers $\bar{e}_a(\bar{\varepsilon}, \tilde{z}, \hat{z})$ in adopting firms, and, therefore, reduces the productivity of adopting firms. If γ goes to infinity, we end up in an equilibrium where the effort effect disappears and all workers in adopting and non-adopting firms provide just the basic effort $\varepsilon = 1$. However, our results regarding incomplete adoption and sorting continue to be valid, as described in corollary 1. In such a situation, adopting firms are only wasting resources from a social perspective since adopting management practices is costly, but they do not have any real productivity effects.

8.3 Social Planner's Problem

The social planner's problem as presented in section 4 implies the following Lagrangian:

$$\left(n_a^*\left(x_a^*\right)^{\frac{\sigma-1}{\sigma}} + n_n^*(q_n^*)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} + \lambda_M(M - n_a^* f_a - n_n^* f_n) +$$
$$\lambda_a\left(L\int_{\tilde{z}^*}^{\infty} e(\bar{\varepsilon}, z)dG(z) - n_a^* x_a^*\right) + \lambda_n\left(L\int_0^{\tilde{z}^*} zdG(z) - n_n^* x_n^*\right).$$

Taking the derivative with respect to n_a^*, n_n^*, x_a^*, x_n^*, and $\tilde{z}^*$ we obtain after modest simplification

$$\frac{\sigma}{\sigma-1}\left(x_a^*\right)^{\frac{\sigma-1}{\sigma}}\left(n_a^*\left(x_a^*\right)^{\frac{\sigma-1}{\sigma}} + n_n^*(q_n^*)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}-1} = \lambda_M f_a + \lambda_a x_a^* \quad (59)$$

$$\frac{\sigma}{\sigma-1}\left(x_n^*\right)^{\frac{\sigma-1}{\sigma}}\left(n_a^*\left(x_a^*\right)^{\frac{\sigma-1}{\sigma}} + n_n^*(q_n^*)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}-1} = \lambda_M f_n + \lambda_n x_n^* \quad (60)$$

$$\left(x_a^*\right)^{-\frac{1}{\sigma}}\left(n_a^*\left(x_a^*\right)^{\frac{\sigma-1}{\sigma}} + n_n^*(q_n^*)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}-1} = \lambda_a \quad (61)$$

$$\left(x_n^*\right)^{-\frac{1}{\sigma}}\left(n_a^*\left(x_a^*\right)^{\frac{\sigma-1}{\sigma}} + n_n^*(q_n^*)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}-1} = \lambda_n \quad (62)$$

$$\frac{\lambda_a}{\lambda_n} = \frac{\tilde{z}^*}{e(\bar{\varepsilon}, \tilde{z}^*)} \quad (63)$$

Using these first order conditions we can now prove the three parts of our proposition 3:
Part (ii): Substituting (61) and (62) into (59) and (60) and then dividing the resulting

expression yields

$$\frac{x_n^*}{x_a^*} = \Phi^\sigma. \tag{64}$$

Equations (20), (21), (22) in the text collectively imply

$$\frac{x_n}{x_a} = \Phi^\sigma.$$

Hence, the social planner would choose relative output levels of n and a varieties that mirror those that the market would generate. Output levels of individual varieties are efficient.

Part (i): The ratio of (61) to (62) combined with (63) and (64) together imply

$$\frac{\tilde{z}^*}{e(\bar{\varepsilon}, \tilde{z}^*)} = \Phi.$$

Comparing this expression to the market equilibrium cutoff shown in equation (22) we have

$$\frac{\tilde{z}^*}{e(\bar{\varepsilon}, \tilde{z}^*)} = \frac{\overline{z}_n(\tilde{z})}{e(\bar{\varepsilon}, \tilde{z})}$$

Because $\overline{z}_n(\tilde{z}) < \tilde{z}$ for all $\tilde{z}$ if must be that $\tilde{z}^* > \tilde{z}$. Too many units of effective labor are allocated by the market to the a technology relative to the output maximizing social planner.

Part (iii): Because the social planner chooses to allocate more labor to the n technology and less to the a technology while choosing the same levels of output of each variety type as in the market equilibrium, it follows directly from the resource constraint that $n_a^* < n_a$ and $n_n^* > n_n$. Since $f_n < f_a$ this means that the market features too few varieties relative to the social optimum.

8.4 Firm-specific Unit Labor Requirements

In section 2.1.2 we abstract from differences in unit labor requirements and normalize α to one. In this section, we discuss firm-type specific unit labor requirements.

Let unit labor requirements of firms of type j be denoted by α_j. Then, equation (22) changes to

$$\Phi_\alpha \equiv \frac{\alpha_n}{\alpha_a}\left(\frac{f_n}{f_a}\right)^{\frac{1}{\sigma-1}}. \tag{65}$$

Coexistence of high- and low-tech firms requires that $\Phi_\alpha < 1$, or

$$\frac{\alpha_n}{\alpha_a} < \left(\frac{f_a}{f_n}\right)^{\frac{1}{\sigma-1}}. \tag{66}$$

Since $f_a > f_n$, $\alpha_n = \alpha_a$ is a sufficient condition for coexistence.

But coexistence is also sustainable if high-tech firms have a lower unit labor requirement. In this case, it is important that the relative difference in unit labor requirements is not too large relative to the difference in fixed costs:

$$(\sigma - 1)(\ln \alpha_n - \ln \alpha_a) \leq (\ln f_a - \ln f_n). \tag{67}$$

One interesting case is where the investment f_j can be interpreted as an innovation in management *and* production technology, where a higher f_j also reduces unit labor requirements α_j. Suppose

$$\alpha_j = f_j^{-\rho}. \tag{68}$$

In this case, coexistence requires that the responsiveness of unit labor requirements with respect to fixed costs is not too large, or

$$\rho < (\sigma - 1)^{-1}. \tag{69}$$

8.5 Multiproduct Firms

Our framework can be extended to make contact with the literature on multiproduct firms and international trade. In the baseline model we focus just on single-product firms, but in this extension, we adopt the flexible manufacturing apparatus present in Eckel and Neary (2010) and demonstrate that multiproduct and single product firms can arise endogenously in our framework. As in Eckel and Neary (2010) and Bernard et al. (2011), trade liberalization induces firms to pare their high-cost product lines that are sold only domestically but the reallocation of labor from small to large firms has the implication that the share of multiproduct firms in total output expands.

In order to extend this framework to multiproduct firms, we follow Eckel and Neary (2010) and assume that all firms possess a certain core competency for a specific variety where their unit labor costs is lowest for all products in their product range. All other products in their product range can then be identified by their (unidimensional) distance to the firm's core competency, denoted by $\omega > 0$. Production of multiple products is subject to flexible manufacturing, which implies that firms can add and drop products to and from their product range freely, but as they add products to their product range and move away from their core competency, unit labor requirements of these products increase. Thus, unit labor requirements α depend on the position ω of a product in a firm's product range, and are increasing in ω:

$$\alpha = \alpha(\omega) \qquad \text{and} \qquad \alpha'(\omega) \equiv \partial\alpha/\partial\omega > 0. \tag{70}$$

To simplify notation we normalize unit labor requirements at the core to one: $\alpha(0) = 1$. In addition, there is a fixed costs per product of f^p.

The profits of a multiproduct firm are

$$\pi_j^{mpf} = \int_0^{\omega_j} Ac_j^{1-\sigma} \alpha(\omega)^{1-\sigma} d\omega - \omega_j f^p - f_j \tag{71}$$

and the optimal product range ω_j is given implicitly by $d\pi_j^{mpf}/d\omega_j = 0$:

$$Ac_j^{1-\sigma} \alpha(\omega_j)^{1-\sigma} = f^p \tag{72}$$

Depending on the size of f^p there are three possible outcomes:

- $f_n < f_a \leq f^p$: All firms are single-product firms
- $f_n \leq f^p < f_a$: Adopting firms are multiproduct firms, non-adopting firms are single-product firms
- $f^p < f_n < f_a$: All firms are multiproduct firms

For this extension suppose that $f^p < f_n < f_a$ so that all firms are multiproduct firms. Combining zero profits with the optimal product range yields

$$\alpha(\omega_j)^{\sigma-1} \int_0^{\omega_j} \alpha(\omega)^{1-\sigma} d\omega - \omega_j = \frac{f_j}{f^p} \tag{73}$$

The left-hand side is increasing in ω_j and the right-hand side is increasing in f_j, so $f_a > f_n$ implies

$$\omega_a > \omega_n \tag{74}$$

Using the same procedure as in section 2.2.4 we obtain the following equation that determines the allocation of labor across firm types:

$$\frac{\bar{z}_n(\tilde{z}_{mpf})}{e(\bar{\varepsilon}, \tilde{z}_{mpf})} = \Phi_{mpf}. \tag{75}$$

Here, Φ_{mpf} depends on the measure of product ranges and is defined as

$$\Phi_{mpf}^{\sigma-1} \equiv \left(\frac{\int_0^{\omega_a} \alpha(\omega)^{1-\sigma} d\omega}{\int_0^{\omega_n} \alpha(\omega)^{1-\sigma} d\omega} \right) \left(\frac{\omega_n f^p + f_n}{\omega_a f^p + f_a} \right) < 1 \tag{76}$$

Since $\omega_a > \omega_n$ and $f_a > f_n$ implies that $\Phi_{mpf} < 1$, $\tilde{z}_{mpf} \in (\underline{z}, \infty)$ and both types of firms coexist. Furthermore, (16) implies that $c_a < c_n$. Thus, the misallocation and the welfare

effects are qualitatively identical. We focused here on the closed economy but extending this extension to an open economy is rather straightforward.

Proposition 13 (Multiproduct Firms) *Adopting firms produce more products.*

Adopting firms pay a lower effective wage rate than non-adopting firms ($c_a < c_n$). This allows them to expand into less efficient activities and produce varieties further away from their core competency with higher unit labor requirements. They have an incentive to do so because SMPs are applicable in all divisions within the firm, so that by adding products to their product range they can lower the fixed costs per product.

8.6 Multiple Export Destinations

Suppose the world is a circle with a continuum of countries, each located at a point on this circle. The distance between countries δ is the shortest arcdistance on the circumference of the circle. Variable trade costs between countries i and j are increasing in distance: $\tau_{ij} = \tau(\delta)$, $\tau(0) = 1$, $\tau'(\delta) > 0$, $\tau''(\delta) > 0$. Note that there is no index on $\tau(\delta)$ so this function is symmetric across all locations. In addition, there are fixed costs f^x per country exported to.

The profits of exporting for firm type j are now given by:

$$\pi_j^{mx} = 2\left(Ac_j^{1-\sigma}\int_0^{\delta_j}\tau(\delta)^{1-\sigma}\,d\delta - \delta_j f^x\right) \tag{77}$$

where δ_j measure the mass of countries firm j exports to. Note that countries are symmetrically exporting to the left and to the right of their location. Hence, profits are multiplied by factor two.

The optimal mass of export destinations is given by

$$\frac{d\pi_j^{mx}}{d\delta_j} = 2\left(Ac_j^{1-\sigma}\tau(\delta_j)^{1-\sigma} - f^x\right) = 0. \tag{78}$$

Since $\tau(0) = 1$ and variable trade costs of exporting to the closest destination are infinitesimally small, our conditions for exporting from section 4 have to be slightly altered:

- $f_n < f_a < f^x$: No exporting
- $f_n < f^x < f_a$: Exporting only by adopting firms
- $f^x < f_n < f_a$: All firms export

For this extension suppose that $f^x < f_n < f_a$ so that all firms export. Aggregate profits for firm of type j are then given by

$$\pi_j^{mx} = A\left(1 + 2\int_0^{\delta_j} \tau\left(\delta\right)^{1-\sigma} d\delta\right) c_j^{1-\sigma} - 2\delta_j f^x - f_j, \tag{79}$$

where mx stands for $m$ultiple e$x$port destinations. Combining free entry and (78) yields

$$\tau\left(\delta_j\right)^{\sigma-1}\left(1 + 2\int_0^{\delta_j} \tau\left(\delta\right)^{1-\sigma} d\delta\right) - 2\delta_j = \frac{f_j}{f^x} \tag{80}$$

Since the left hand side of (80) is increasing in δ_j and the right hand side is increasing in f_j, $f_a > f_n$ implies

$$\delta_a > \delta_n. \tag{81}$$

Using the same procedure as in section 2.2.4 we obtain the following equation that determines the allocation of labor across firm types:

$$\frac{\bar{z}_n\left(\tilde{z}_{mx}\right)}{e\left(\bar{\varepsilon}, \tilde{z}_{mx}\right)} = \Phi_{mx}. \tag{82}$$

Here, Φ_{mx} depends on the measure of export destinations and is defined as

$$\Phi_{mx}^{\sigma-1} \equiv \left(\frac{1 + 2\int_0^{\delta_a} \tau\left(\delta\right)^{1-\sigma} d\delta}{1 + 2\int_0^{\delta_n} \tau\left(\delta\right)^{1-\sigma} d\delta}\right)\left(\frac{2\delta_n f^x + f_n}{2\delta_a f^x + f_a}\right) < 1 \tag{83}$$

Since $\delta_a > \delta_n$ implies that $\Phi_{mx} < 1$, $\tilde{z}_{mx} \in (\underline{z}, \infty)$ and both types of firms coexist. Furthermore, (16) implies that $c_a < c_n$. Thus, the misallocation and the welfare effects are qualitatively identical.

Proposition 14 (Multiple Export Destinations) *Adopting firms export to more destinations and ship longer distances.*

8.7 Multiple Degrees of Adoption

In this extension we illustrate that our assumption of just two degrees of adoption, adopt and non-adopt, and consequently just two types of firms, is not critical for the mechanism we describe. The simple case with just two types of firms is very useful for understanding the forces behind the mechanism, and for providing intuition, but it needs to be extended for empirical work. Here we show a straightforward way to extend this framework to include multiple degrees of adoption that lead to a higher degree of firm heterogeneity. We can think

of these as different degrees of adopting SMPs.

Suppose there are N degrees of adopting SMPs $j \in \{0, ..., N\}$, and each degree allows firms to observe workers up to a specific skill level. A degree j allows firms to observe up to skill level $\breve{z}_j$, so that all skills with $z < \breve{z}_j$ can be observed and all skills with $z > \breve{z}_j$ cannot. This also implies that only workers with skills $z < \breve{z}_j$ can be incentivized, and workers with skills $z > \breve{z}_j$ cannot. A higher degree of adoption (with a higher $\breve{z}_j$) is more expensive and requires strictly higher fixed costs, so that $f_{j+1} > f_j > 0$.[29]

Given that a firm with adoption level j (type-j firm) can only observe skills up to level $\breve{z}_j$, it can only pay a piece rate up to this level. Above this level, it has to pay a wage based on the average productivity.[30] Hence, the wage paid by a type-j firm to workers with skill z depends on the skill of these workers:

$$w_j(z) = c_j \times \begin{cases} e(\bar{\varepsilon}, z) & \text{for} \quad \tilde{z}_j < z < \breve{z}_j < \tilde{z}_{j+1} \\ \bar{e}(1, \breve{z}_j, \tilde{z}_{j+1}) & \text{for} \quad \tilde{z}_j < \breve{z}_j < z < \tilde{z}_{j+1} \end{cases}, \tag{84}$$

where $\bar{e}(\varepsilon, z_1, z_2) \equiv \int_{z_1}^{z_2} e(\varepsilon, z)\, dG(z) / [G(z_2) - G(z_1)]$. The set of workers that sort into the labor pool of type j is $Z_j = \{z : z \in (\tilde{z}_j, \tilde{z}_{j+1})\}$, so that $\tilde{z}_j$ and $\tilde{z}_{j+1}$ denote the lower and upper boundary of skills in firms of type j.

[Figure 5 here]

The profile of wages in type-j firms is illustrated in Figure 5. In the range between $\tilde{z}_j$ and $\breve{z}_j$, firms of type j can fully observe the true productivity of individual workers and pay a wage based on this productivity: $w_j(z) = c_j e(\bar{\varepsilon}, z)$. Above $\breve{z}_j$, these firms only know the average productivity of their workforce because their degree of adoption of SMPs does not allow them to monitor workers with these skills. As a consequence, workers are also not incentivized, and the wage is based on $\bar{e}$: $w_j(z) = c_j \bar{e}(1, \breve{z}_j, \tilde{z}_{j+1})$. All workers in a type-$j$ labor pool receive the same effective wage rate c_j.

Note that there is a discrete jump in the wage profile within a labor pool. This jump is due to the fact that the average skills of workers in the range $(\breve{z}_j, \tilde{z}_{j+1})$ is discretely higher than the skills of worker $\breve{z}_j$. However, the jump is mitigated by the fact that workers in the range $(\breve{z}_j, \tilde{z}_{j+1})$ are not incentivized and, therefore, only provide the minimal effort $\varepsilon = 1$.

The boundaries of a particular labor pool are determined by the sorting conditions. In the case of $N + 1$ technologies there are N sorting conditions. For the boundary between

[29] Our modelling with many screening technologies is similar to a discrete version of screening in Helpman, Itskhoki and Redding (2010).

[30] We maintain the assumption that firms know the distribution of skills in their labor pool.

firms of type j and $j-1$, this sorting condition is

$$\frac{c_j}{c_{j-1}} = \frac{\bar{e}\left(1, \breve{z}_{j-1}, \tilde{z}_j\right)}{e\left(\bar{\varepsilon}, \tilde{z}_j\right)} \tag{85}$$

If this condition is fulfilled, workers with skill $\tilde{z}_j$ receive the same wage in the two labor pools j and $j-1$. Since $\tilde{z}_j > \breve{z}_{j-1}$, their skill is above the observable skill in type-$(j-1)$ firms, so their wage in the type-$(j-1)$ labor pool is $c_{j-1}\bar{e}\left(1, \breve{z}_{j-1}, \tilde{z}_j\right)$. In the type-$j$ labor pool their skill is below the threshold of observability, so firms with adoption degree j will pay them a wage based on their true productivity of $c_j e\left(\bar{\varepsilon}, \tilde{z}_j\right)$. We assume that condition (17) holds locally.

Given this sorting condition, the marginal skill $\tilde{z}_j$ is then determined by

$$\frac{\bar{e}\left(1, \breve{z}_{j-1}, \tilde{z}_j\right)}{e\left(\bar{\varepsilon}, \tilde{z}_j\right)} = \Phi_j, \tag{86}$$

where

$$\Phi_j^{\sigma-1} \equiv \frac{f_{j-1}}{f_j} < 1, \tag{87}$$

and our key corollaries with respect to relative revenues per worker (corollary 3) and average wages (corollary 4) also hold in this extension.

Proposition 15 (Discrete Adoption) *With multiple degrees of adoption, sorting leads to multiple labor pools where different degrees of adoption can coexist. Firms with a higher degree of adoption (a higher observable skill level $\breve{z}$) recruit more productive workers and exhibit higher revenues per worker and higher average wages.*

The market equilibrium in the case with multiple degrees of adoption is similarly distorted as the equilibrium with only two levels of adoption. Because wages of workers with above average skills in each labor pool receive wages based on their average productivity and not on their marginal productivity, the skill set of workers sorting into a particular labor pool ($j > 0$) is too low compared to the social optimum ($\tilde{z}_j < \tilde{z}_j^*$). As a consequence, for any degree of adoption $j > 0$, employment in firms with a lower degree of adoption is too low and in firms with a higher degree of adoption too high compared to the social optimum: $G\left(\tilde{z}_j\right) L < G\left(\tilde{z}_j^*\right) L$ and $\left[1 - G\left(\tilde{z}_j\right)\right] L > \left[1 - G\left(\tilde{z}_j^*\right)\right] L$.

Finally, we want to point out that the assumption of a discrete number of degrees of adoption is important. If N goes to infinity, each labor pool becomes infinitesimal, and the boundaries of the pools collapse onto their observable skills. Formally, $\tilde{z}_j$ approaches $\breve{z}_j$ from

below, and $\tilde{z}_{j+1}$ approaches $\breve{z}_j$ from above. Then,

$$\lim_{\tilde{z}_{j+1} \to \breve{z}_j} \bar{e}_j \left(1, \breve{z}_j, \tilde{z}_{j+1}\right) = e_j \left(1, \breve{z}_j\right), \tag{88}$$

and there is no more pooling in the labor market and no more uncertainty about skill levels. Each firm gets exactly the skill level that they target with their technology, and effective wages are equalized across different degrees of adoption. As a consequence, the inefficiency associated with pooling also disappears.

Proposition 16 (Continuous Technologies) *With continuous adoption possibilities, there is no more pooling in the labor market and the inefficiency associated with pooling disappears.*

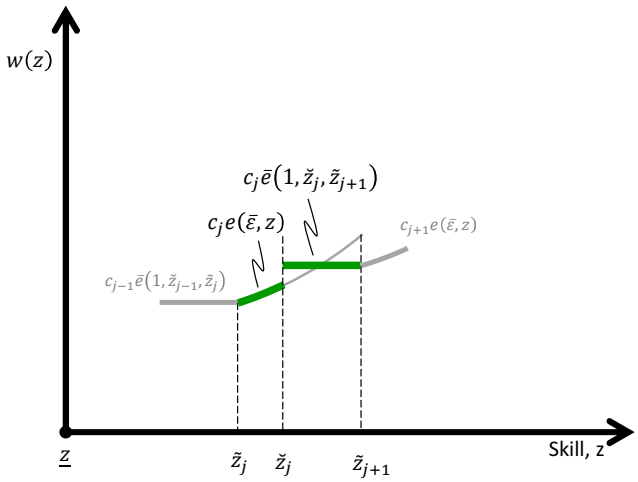

Figure 5: Multiple Screening Technologies