

Fragmentation of Production and the Wage Distribution

Elisa Rubbo

May 8, 2023

Abstract

This paper proposes a novel partial equilibrium theory of the fragmentation of production across industries in an environment where firms have variable input requirements over time, while workers have different comparative advantage over inputs and cannot be freely reallocated across firms. Fragmenting production allows firms to allocate workers to tasks based on their comparative advantage at all times, but it requires costly trading of inputs between firms. I show that, in general equilibrium, there is a two-way relationship between technical change, fragmentation, and the wage distribution. On the one hand, reducing the cost of trading inputs between firms leads to a hollowing out of the wage distribution. On the other hand, an increase in the skill premium strengthens the incentives to fragment production. The model illustrates how skill biased technical change not only affects the wage distribution, but also the organization of production. Additionally, the theory suggests that outsourcing – which is a form of fragmentation – could be the consequence as well as the cause of a higher skill premium.

1 Introduction

In this paper I propose a novel partial equilibrium theory of the fragmentation of production across industries, designed to highlight the interconnection between technical change, fragmentation, and the wage distribution in general equilibrium. I consider an environment where firms have variable input requirements over time, heterogeneous workers have different comparative advantage over inputs, and there are frictions to reallocating workers

across firms. Fragmenting production allows workers to specialize on the inputs where they have comparative advantage, but it entails a productivity loss in the form of trade costs between firms. In partial equilibrium, this tradeoff pins down the division of production across industries. In general equilibrium, changes in technology and trade costs affect the workers' marginal product and the extent to which they can specialize, and thus also their wages. As a result there is a two-way relationship between the wage distribution and the fragmentation of production across industries. Under realistic assumptions I show that a decline in fragmentation costs increases the skill premium, and leads to a hollowing out of the wage distribution. On the other hand, a rise in the skill premium strengthens the incentives to fragment production.

The model suggests that seemingly different empirical facts are interconnected. The skill premium has increased substantially over the past decades (Katz and Murphy 1992; Katz and Autor 1999), widening the wage gap between top earners and the other workers. Middle income workers were especially hurt, a phenomenon that Autor and Dorn (2013) refer to as the hollowing out of the wage distribution. At the same time, high- and low-wage workers are more and more segregated across firms (Song et al. (2019)), and outsourcing of low-skill services is on the rise (Goldschmidt and Schmieder (2017)). In my framework outsourcing and fragmentation are the same phenomenon, and I will use the two terms interchangeably. My model implies that both skill-biased technical change and a lower cost of trading inputs between firms increase outsourcing and the skill premium at the same time, thus replicating their joint behavior in the data.

My theory highlights that biased technical change has implications for the organization of production, not just for the wage distribution. It also shows that changes in fragmentation costs – even when they are uniform and thus unbiased ex-ante – are a form of biased technical change ex-post. The model rationalizes outsourcing of both high- and low-skill services, consistent with the data. Instead, the recent literature which views outsourcing as a way to avoid sharing “firm premia” can only rationalize outsourcing of low-skill services. Moreover, according to the “firm premium” theory, outsourcing is the cause of changes in the wage distribution. I raise the caveat that it may also be their consequence.

Section 2 illustrates the setup. I move from a standard production function, with multiple intermediate inputs and final goods. Producers have constant desired markups, offset by a combination of profit taxes and input subsidies. Absent any friction, the division of production across industries would be indeterminate. That is,

whether intermediate inputs are produced by final good firms or by separate input suppliers is irrelevant for profits and welfare. With volatile input needs and frictional worker reallocation, instead, the division of production between final good firms and input suppliers is uniquely pinned down.

The key idea is that firms need a different input mix at each point in time. For example, they do not launch new products every day, and so they need designers and marketing experts only rarely. Similarly, their equipment only breaks down every so often, thus they do not always need technicians to repair it. At the same time, different workers have a comparative advantage in making different inputs. Therefore, in a frictionless world, firms would like to adjust their workforce composition instant by instant based on their needs. However this is not feasible, because firms cannot freely exchange workers between each other. In the model, firms must hire workers through long term contracts that are renewed at discrete time intervals ($t, t + 1$, etc.). Production instead happens throughout a continuous time period ($t, t + 1$), and at each instant $r \in (t, t + 1)$ final good firms need a different set of intermediate inputs. As a result, input needs change more frequently than firms can adjust their in-house skill composition, so that in-house workers are idle for part of the time, or employed on tasks where they do not have comparative advantage.

To mitigate this friction, final good firms can purchase inputs on the spot market. This applies to both physical inputs and labor services. For example, when they want to launch a new product, firms can hire design engineers or marketing experts from a consulting company, or if they need to repair equipment they can hire technicians from a maintenance company. Final good firms can adjust the input mix that they purchase on the spot market instant by instant, but they face a trade cost (a melting iceberg cost in the model), which could represent transportation costs, lack of customization, etc. Spot-market input providers in turn hire workers to produce, and their workers specialize on a single input (i.e. they work on that input for the whole time).

To simplify the analysis I assume that the demand for inputs is i.i.d. across firms at each point in time. Therefore the overall industry-level demand for each given input is constant over time, and thus spot-market producers can offer perfect insurance. I also rule out economies of scale, by assuming that input needs are perfectly correlated across all units of the final product within a given firm. Likewise I rule out economies of scope, by assuming that sharing inputs between different products within a firm triggers the same iceberg cost as purchasing inputs on the spot market. Accounting for economies of scale and scope, and for limited insurance through the spot market,

would deliver richer partial-equilibrium patterns of fragmentation across the firm size distribution. However the interconnections between fragmentation and the wage distribution in general equilibrium would be the same.

Section 3 analyzes the equilibrium of the model, highlighting some key properties. First, I show that the market equilibrium is efficient. I then study the optimal allocation of workers to firms, and the optimal division of production across firms. Firms can decide to produce each input entirely in-house, to purchase it entirely on the spot market, or to produce some units in house and purchase the remaining ones on the spot market as needed. I show that firms use the spot market more intensively for inputs that are needed with low probability relative to the trade cost. If there are no trade costs, the spot market is always (weakly) preferable. Moreover, if all workers had the same relative productivities on the various inputs there would be no role for the spot market. Otherwise, I show that workers who have a strong comparative advantage on one input are employed by spot market producers, while workers who have similar productivities on multiple inputs are kept in-house by final producers. This is intuitive, because it is more costly for final producers to employ a worker with strong comparative advantage on a suboptimal input.

The properties just outlined hold in a general setting. To derive clean comparative statics for spot market trade shares and relative wages, in Sections 4 and 5 I restrict attention to two special cases which impose more structure on the skill distribution. The skill distribution determines the supply elasticities of the various inputs, and therefore it plays a key role in shaping the response of wages to trade costs and biased productivity changes.

To reproduce the hollowing out of the wage distribution, in Section 4 I assume that there is an ordering of inputs such that high-skill workers have a comparative advantage on high-index inputs (i.e. there is supermodularity between skills and tasks). I also assume that the skill density is concentrated in the low- to mid-skill region. The supermodularity assumption implies that workers at the top and bottom ends of the skill distribution have a strong comparative advantage on the most and the least skill-intensive tasks respectively, while workers in the middle have similar skills on a range of tasks. Therefore the workers at the extremes of the wage distribution are employed by spot market producers, and specialize on the tasks where they have comparative advantage. Mid-skill workers instead are employed by final good producers, and perform a variety of tasks.

Suppose now that the trade cost that firms face when using the spot market declines uniformly for all inputs. While unbiased ex-ante, this technical change ends up being biased in favor of the highest and lowest income workers,

because these workers are employed by spot market producers and thus can reap the productivity benefits. As it is well known, a biased productivity change does not necessarily increase the relative wage of the workers who become more productive, because higher productivity is counterbalanced by higher supply and therefore lower prices. The productivity effect dominates when input demand is price-elastic, and when initial outsourcing shares are small, so that the supply effect is also small. Indeed, when initial outsourcing shares are small enough, the relative wage of outsourced workers increases even when inputs are complementary in production. As a result the workers in the middle of the wage distribution lose relative to those at the extremes. Moreover, when there are fewer top-skill than low-skill workers, a decline in trade costs raises the overall supply of low-skill inputs relative to high-skill ones, which in turn increases the skill premium.¹

While Section 4 studies the effect of outsourcing costs on the wage distribution, Section 5 studies the effect of skill-biased technical change on outsourcing. Intuitively, as the skill premium increases it becomes more and more costly for firms to employ top-skill workers on less skill-intensive tasks where they do not have comparative advantage. Therefore it is optimal to fragment the production of high- and low-skill inputs across different industries.

To make the model more concrete, and to highlight how it can rationalize low-skill outsourcing, I give a specific interpretation to the suboptimal tasks that top-skill workers may end up performing when hired in house. That is, top-skill workers might occasionally have to hire and supervise low-skill workers. This task is outside of their comparative advantage, and ideally it should be performed by mid-skill human resource (HR) managers. As the skill premium increases, it becomes more and more costly for final good firms to use top-skill workers as HR managers, and therefore they start to outsource the HR management tasks. If low-skill workers must be hired by the same firm as their managers, outsourcing the manager also implies outsourcing the worker. This form of outsourcing per-se has no consequence for the allocation of low-skill workers to final good firms, nor directly for their wages. Nonetheless, we observe that the outsourcing share and the skill premium increase at the same time.

A natural reason why the skill premium may increase is skill-biased technical change. I represent this as an increase in the productivity of the high-skill service sector, and show that the model, calibrated to conventional parameter values, matches well the joint evolution of the skill premium, the outsourcing share in manufacturing, and the overall sale and employment share of manufacturing.

¹As a caveat, this argument is valid only for small initial outsourcing shares. If inputs are complementary, further declines in outsourcing costs would have an opposite effect on the wage distribution.

1.1 Related literature

Wernerfelt (2015) develops a rich partial equilibrium model of the tradeoff between keeping idle resources in house and purchasing them on a spot market, subject to a frictional cost. A key contribution of my work is to embed this tradeoff in a general equilibrium model, and to study the implications for the wage distribution. Garicano and Rossi-Hansberg (2006) also study the general equilibrium relation between trade costs (communication costs in their language), the organization of production, and the wage distribution, based on a model of endogenous sorting and matching of workers with a one-dimensional skill distribution. They find that lower communication costs generate a superstar (Rosen (1981)) effect, by allowing top-skill workers to specialize in harder tasks while being matched to a larger number of low-skill workers. I interpret skill more broadly as a multi-dimensional object, which allows me to obtain non-monotonic effects along the wage distribution, and to match a richer set of fragmentation patterns beyond tree-shaped networks.

Hart and Moore (1990) elaborated a theory of firm boundaries based on the optimal choice between arms-length trade and integration. While Hart and Moore (1990) restrict attention to transactions between a buyer and a single supplier, I study the optimal choice between purchasing from a single vs multiple suppliers (i.e. producing in-house vs using the spot-market). Relatedly, Costinot et al. (2013) and Kikuchi et al. (2018) model the division of production along vertical chains. My model can rationalize richer fragmentation patterns, not just vertical chains. I also extend the analysis to study the connection between fragmentation and wages.

A large literature seeks to explain the rise in the skill premium. See Goldin and Katz (2010), Acemoglu and Autor (2011) and Acemoglu and Restrepo (2018) for theories that elucidate the relationship between technical change and the wage distribution, but do not deliver implications for the organization of production. Eckert et al. (2022) underscore the role of high-skill service industries and the geographical concentration of top earners around large cities, a point also made by Kleinman (2023). A large literature (Grossman and Helpman (2018); Grossman et al. (2017); Grossman and Rossi-Hansberg (2008); Feenstra and Hanson (1996)) relates changes in the wage distribution with international trade, tracing them back to a different skill composition across countries. My focus is on trade costs between industries within the same country, assuming a constant skill composition over time.

A recent literature connects the skill premium with outsourcing of low skill workers (Goldschmidt and Schmierer (2017); Bilal and Lhuillier (2020)), which allows more productive firms to avoid sharing with them a “firm premium”.

While capturing an important incentive behind low-skill outsourcing, this explanation cannot rationalize outsourcing of high-skill workers, nor can it account for a larger wage decline in the middle than at the bottom of the distribution. Moreover, according to this explanation causality runs from outsourcing to the wage distribution, while I argue that both directions are possible.

2 Model

This section lays out the assumptions about preferences and production functions, and introduces our model of the tradeoff between producing inputs in house or buying them on the spot market. It then illustrates the solution to the firms' cost minimization problem, and introduces the equilibrium definition.

In the model, time is divided into discrete intervals $[t, t + 1]$, and we denote each instant between t and $t + 1$ by $r \in (t, t + 1)$. Decisions within one period $[t, t + 1]$ are independent from past and future periods, therefore we focus on a single period, omitting time subscripts whenever this doesn't cause confusion.

2.1 Consumers-workers

There is a unit mass of agents in the economy, all with identical per-period consumption preferences

$$\mathcal{C}(c_1, \dots, c_N)$$

over the N final goods produced in the economy.

Consumers are endowed with one unit of labor at each point in time $r \in (t, t + 1)$, and they are characterized by a skill vector \mathbf{s} (described below) which determines their wage $W(\mathbf{s})$. We denote by \mathcal{S} the set of skills, and by μ the density of agents over $\mathbf{s} \in \mathcal{S}$. The total mass of workers is \mathcal{M} .

Wages are paid and consumption decisions are made at the end of each time interval.

There is no borrowing and lending between agents, and our assumptions will imply that there are no profit rebates

(see the production section below). Hence the budget constraint for an agent with skill \mathbf{s} is given by

$$\sum_i P_i c_i(\mathbf{s}) = W(\mathbf{s})$$

Consumers choose the consumption bundle (c_1, \dots, c_N) that maximizes utility given prices and the budget constraint. We assume that the utility function \mathcal{C} is homothetic, and we denote the associated price index by P_C .

2.2 Production

Final goods There are N final goods in the economy, indexed by $i \in \{1, \dots, N\}$. For each final good there is a fixed unit mass of varieties (or blueprints). Consumers have CES preferences over varieties, so that the total output of final good i is given by (again omitting time subscripts for legibility):

$$Y_i = \left[\int_0^1 Y_{if}^{\frac{\epsilon-1}{\epsilon}} df \right]^{\frac{\epsilon}{\epsilon-1}}$$

where Y_{if} is the output of variety f and ϵ is the elasticity of substitution between varieties.

To simplify the analysis and ensure efficiency, I make the following assumption.

Assumption 1. *Markups are exactly offset by input subsidies, financed through lump-sum taxes on profits. On net, taxes and profit rebates cancel out.*

All varieties of final good i have the same constant-returns-to-scale production function G_i , which takes as inputs K intermediate goods (indexed by $k \in \{1, \dots, K\}$):

$$Y_{if} = G_i(A_{i1}X_{i1f}, \dots, A_{iK}X_{iKf})$$

where X_{ikf} is the quantity of intermediate good k used by firm f in industry i . Inputs are assembled into final goods at the end of each time period.

Intermediate inputs Intermediate inputs are produced throughout the whole interval $(t, t + 1)$, using labor. Each agent is endowed with a vector of skills $\mathbf{s} = (s_1, \dots, s_K)$, which corresponds to their productivity on each of the intermediate inputs.

Final good firms can hire workers directly and produce inputs in house, or they can buy inputs on the spot market. Spot market input suppliers also use labor to produce. The market for intermediate inputs is perfectly competitive, so that the number of firms on this market is indeterminate.

To model volatility in the input needs of final producers, we assume that input k is needed by firm f in industry i only for a subset of the instants $r \in (t, t + 1)$. At the end of the period, the total quantity of input k available to firm f is

$$X_{ikf} = \int_t^{t+1} \mathbb{I}_{ikf}(r) Q_{ikf}(r) dr$$

where $Q_{ikf}(r)$ is the quantity of input k available to the firm at time r , while $\mathbb{I}_{ikf}(r) = 1$ if input k is needed at time r , and $\mathbb{I}_{ikf}(r) = 0$ otherwise. In turn, the quantity $Q_{ikf}(r)$ of input k available to firm f at time r is given by

$$Q_{ikf}(r) = \int_{\mathcal{S}} s_k \ell_{ikf}(s, r) ds + \tau_k X_{ikf}^{\mathcal{O}}(r)$$

where $s_k \ell_{ikf}(s, r)$ are the efficiency units of labor that firm f is employing on input k in house, $X_{ikf}^{\mathcal{O}}(r)$ is the quantity of input k that the firm is purchasing on the spot market, and τ_k is an iceberg trade cost applied to input k whenever it is purchased on the spot market. Note that in general $\ell_{ikf}(s, r)$ will not constant be over time r , because final good producers optimally reallocate in-house workers across tasks, based on their needs.

Analogously, at each point in time r , spot-market provider f of input k has output

$$X_{kf}(r) = \int_{\mathcal{S}} s_k \ell_{kf}(s, r) ds$$

To model the tradeoff between producing in house and using the spot market, we make the following key assumption.

Assumption 2. *Both final good firms and intermediate input producers must hire workers at time t for the whole time period $(t, t + 1)$. Therefore firms cannot adjust their workforce composition based on the set of inputs they need at each point in time $r \in (t, t + 1)$. However, firms can adjust the quantity of inputs that they buy on the spot*

market at any point in time.

Assumption 2 could represent relationship contracts, or labor market regulation which prevents firms from hiring and firing workers at high frequencies. In this context, the spot market is a way to resolve the tension between the frequency at which needs change, and the frequency at which firms can adjust their workforce.

Before turning to the firms' optimal choice of which inputs to make or buy, and which skills to hire, I introduce some additional simplifying assumptions.

Assumption 3. *Input needs are iid across firms f , industries i , and instants r .*

Assumption 4. *Input needs are perfectly correlated across all units of a given variety.*

Assumption 5. *Whenever the same workers produce an intermediate input for at least two final good varieties, all units of the input are subject to the full iceberg cost.*

By Assumption ?? the probability $\mathbb{P}(\mathbb{I}_{ikf} = 1)$ is independent of the time r and the identity of the firm f , while by Assumption ?? the probability $\mathbb{P}(\mathbb{I}_{ikf} = 1)$ is independent of total output Y_{if} . We denote by $\pi_{ik} \equiv \mathbb{P}(\mathbb{I}_{ikf} = 1)$.

Remark 1. Note that, as a consequence of Assumption 3, the demand for intermediate input k is constant over time. Indeed, Assumption 3 guarantees that firms can perfectly insure against their input volatility by using the spot market – subject to the iceberg cost.

Remark 2. Assumption 4 rules out economies of scale, while Assumption 3 rules out economies of scope. By Assumption 4, larger firms have equally volatile needs as small firms. By Assumption 5, multi-product firms incur the same productivity loss when they reallocate idle workers across their own varieties as when they purchase labor services on the spot market.

Summary of timing Table 1 provides a schematic summary of the actions taken at each instant $r \in [t, t + 1]$.

Time t	All producers f (of final goods and intermediate inputs) decide how many workers $L_f(s)$ to hire for each skill type s , taking as given the wage profile $W(s)$.
----------	---

Time $r \in (t, t+1)$	Final good producers optimally allocate their in-house workers to producing the inputs that they need, and they decide how much of each input to purchase on the spot market. Spot market suppliers also produce and sell inputs.
-----------------------	--

Time $t+1$	Final good producers assemble inputs into final goods and sell final goods to consumers. Wages are paid and consumption takes place.
------------	---

Table 1: Summary of timing

Cost minimization problem All producers take the wage distribution as given. Spot market producers of input k hire worker types

$$\mathbf{s} \in \operatorname{argmin}_{\mathbf{s}'} \frac{W(\mathbf{s}')}{s'_k} \quad (1)$$

and charge prices

$$P_k = \frac{W(\mathbf{s})}{s_k} \quad (2)$$

Denoting by $L_k(\mathbf{s})$ the number of workers of skill \mathbf{s} employed by spot market producers of k , the total quantity of input k on the spot market is

$$X_k^{\mathcal{O}} = \int_{\mathcal{S}} s_k L_k(\mathbf{s}) ds$$

Thanks to constant returns and Assumptions 4 and 5, all final good producers within an industry face the same marginal cost and choose the same input combination. Omitting the firm subscript f for legibility, the cost minimization problem for industry i is:

$$\min_{L_i(\cdot), \ell_{ik}(\cdot, \cdot), Q_{ik}^{\mathcal{O}}(\cdot)} \int W(s) L_i(s) ds + \sum_k \int_t^{t+1} P_k X_{ik}^{\mathcal{O}}(r) dr \quad (3)$$

subject to the constraints

$$\begin{aligned}
1 &= G_i (A_{i1}X_{i1}, \dots, A_{iK}X_{iK}) \\
X_{ik} &= \int_t^{t+1} \mathbb{I}_{ik}(r) \left[\int_{\mathcal{S}} s_k \ell_{ik}(s, r) ds + \tau_k X_{ik}^{\mathcal{O}}(r) \right] dr \quad \forall k \\
L_i(s) &= \sum_k \ell_{ik}(s, r) \quad \forall r, s
\end{aligned}$$

With constant returns to scale the multiplier on the first constraint is equal to the marginal cost, and, given Assumption 1, this is also equal to the price P_i of final good i . From the first order condition (??) below, the multipliers on the second set of constraints are equal to the marginal revenue products of each input i in industry k . We denote these multipliers by W_{ik} , and refer to them as the “wage” of task k in industry i . Finally, we denote the multipliers on the third set of constraints by $\eta(r, s)$.

The first order conditions of problem (3) are

$$W_{ik} = P_i A_{ik} G'_{ik} \tag{4}$$

$$W(s) \geq \int_0^1 \eta(r, s) dr, \quad \left(W(s) - \int_0^1 \eta(r, s) dr \right) L_i(s) = 0 \quad \forall s \tag{5}$$

$$\eta(r, s) \leq \mathbb{I}_{ik}(r) W_{ik} s_k \tag{6}$$

$$W_{ik} \leq \frac{P_k}{\tau_k}, \quad \left(W_{ik} - \frac{P_k}{\tau_k} \right) \left[\int_0^1 X_{ik}^{\mathcal{O}}(r) dr \right] = 0 \quad \forall k \tag{7}$$

Condition (6) implies that, at each point in time r , every skill type s is assigned to a task

$$k \in \arg \max_{k'} W_{ik'} s_{k'} \mathbb{I}_{ik'}(r)$$

so that

$$\eta(r, s) = \max_k W_{ik} s_k \mathbb{I}_{ik}(r)$$

Thus, for each industry-type pair (i, \mathbf{s}) , there is a pecking order of tasks $k_1(i, \mathbf{s}), \dots, k_K(i, \mathbf{s})$ characterized by

$$W_{ik_1} s_{k_1} \geq \dots \geq W_{ik_K} s_{k_K}$$

and, at each instant r , the worker is assigned to the first task in the Pecking order that is needed. The probability that worker type \mathbf{s} is assigned to the h -th task in the pecking order is given by

$$\mathbb{P}(s \text{ on } k_h \mid s \text{ hired by } i) = \pi_{ik_h} \prod_{h' < h} (1 - \pi_{ik_{h'}})$$

We can then rewrite condition (5) as

$$W(s) \geq \sum_k W_{ik} s_k \mathbb{P}(s \text{ on } k \mid s \text{ hired by } i) \tag{8}$$

and final producers in industry i only hire skill types \mathbf{s} such that

$$W(s) = \sum_k W_{ik} s_k \mathbb{P}(s \text{ on } k \mid s \text{ hired by } i)$$

Moreover, by condition (7), final producers in industry i purchase a positive amount of input k on the spot market only if

$$W_{ik} = \frac{P_k}{\tau_k}$$

2.3 Summary of notation

Table 2 below summarizes our notation.

Output of final goods	$\{Y_i\}, i = 1, \dots, N$
Quantity of intermediate inputs	$\{X_{ik}\}, i = 1, \dots, N, k = 1, \dots, K$
Prices of final goods	$\{P_i\}, i = 1, \dots, N$
Input prices on the spot market	$\{P_k\}, k = 1, \dots, K$
Marginal revenue product of inputs	$W_{ik} \leq P_k \frac{\pi_{ik}}{\tau_k}$
Skill vectors	$\mathbf{s} = (s_1, \dots, s_K) \in \mathcal{S} \subseteq \mathbb{R}^K$
Density of workers over skills	$\mu \in \Delta(\mathcal{S})$
Instantaneous probability of needing input k	$\pi_{ik} \equiv \mathbb{P}(\mathbb{I}_{ik} = 1)$
Outsourcing costs	$\{\tau_k\}, i = 1, \dots, N, k = 1, \dots, K$
Input-by-sector productivities	$\{A_{ik}\}, i = 1, \dots, N, k = 1, \dots, K$

Table 2: Summary of Notation

2.4 Equilibrium

For given productivities $\{A_{ik}\}$, input need probabilities $\{\pi_{ik}\}$, and trade costs $\{\tau_k\}$, the general equilibrium is given by a wage function $W(s) : \mathcal{S} \mapsto \mathbb{R}$, a set of good prices $\{P_i\}, \{P_k\}$ and marginal revenue products $\{W_{ik}\}$, hiring functions $\{L_i(s)\}$, and allocation functions $\{\ell_{ik}(s, r)\}$ such that consumers maximize utility subject to their budget constraints, conditions (1), (2), (4)-(6) are satisfied, and the following market clearing conditions hold:

$$\sum_i L_i(s) = \mathcal{M}\mu(s) \quad \forall s \quad (9)$$

$$\sum_i X_{ik}^\mathcal{O}(r) = X_k^\mathcal{O} \quad \forall r, k \quad (10)$$

Proposition 1. *An equilibrium always exists and is unique. Under Assumption 1, the equilibrium is efficient.*

Proof. For existence and uniqueness, see the Appendix. For efficiency, it is immediate to verify that the optimality conditions for the social planner are implied by the consumers' first order conditions together with the producers' optimality conditions (1), (2), and (4)-(6), plus the market clearing conditions (9) and (10). \square

3 General results

In this section I state some general properties of the equilibrium allocation of workers, and derive comparative statics for input prices with respect to productivity and trade costs.

3.1 Allocation of workers and wage distribution

Let's begin by defining the set \mathcal{Z} of all possible worker allocations.

Definition 1. An allocation $\zeta \in \mathcal{Z}$ is a pair of an industry and a pecking order of tasks, such that the worker is employed on the first task whenever it is needed, and otherwise she is employed on the second task, and so on. We denote by $i(\zeta)$ the industry corresponding to allocation ζ , and by $\mathbb{P}_\zeta(k)$ the probability that the worker is employed on task k given the pecking order implied by ζ . Finally, we define a set of probability distributions $\{\xi(\mathbf{s}, \zeta)\}_{\mathbf{s} \in \mathcal{S}} \in \Delta(\mathcal{Z})$ such that $\xi(\mathbf{s}, \zeta)$ is the fraction of workers with skill \mathbf{s} that are allocated according to $\zeta \in \mathcal{Z}$ in equilibrium.

Note that the set \mathcal{Z} is finite, with $N \times 2^K + K$ elements, corresponding to all possible pecking orders of tasks in each final industry plus employment by input providers.

We include spot market producers among the set of industries i , so that $i \in \{1, \dots, N + K\}$, and – with some abuse of notation – we denote $W_{ik} = P_k$ when i is the spot market producer of input k . Conditions (4) and (2) imply that $\xi(\mathbf{s}, \zeta) > 0$ only if

$$\sum_{k=1}^K W_{i(\zeta)k} \mathbb{P}_\zeta(k) s_k \in \max_{\zeta' \in \mathcal{Z}} \sum_{k=1}^K W_{i(\zeta')k} \mathbb{P}_{\zeta'}(k) s_k \quad (11)$$

Note that the maximum on the right hand side of equation (11) needs not be unique, therefore for the same \mathbf{s} we can have $\xi(\mathbf{s}, \zeta) \in (0, 1)$ for multiple allocations ζ (we will see an example of this in Section 5 below).

Given equation (11) and the first order conditions (2) and (4), equilibrium wages are given by

$$W(\mathbf{s}) = \sum_{\zeta \in \mathcal{Z}} \xi(\mathbf{s}, \zeta) \sum_{k=1}^K W_{i(\zeta)k} \mathbb{P}_\zeta(k) s_k \quad (12)$$

Building on condition (11), the next Lemmas highlight some intuitive properties of the equilibrium allocation. First, Lemma 1 provides a threshold for the outsourcing cost relative to the volatility of input needs below which final producers do not find it profitable to outsource. In particular, firms only produce inputs in house when $\{\tau_k\}_{k=1}^K = 0$.

Lemma 1. *If $\tau_k < \pi_{ik}$, then $X_{ik}^{\mathcal{O}}(r) = 0 \forall r \in (t, t + 1)$*

Proof. By contradiction, if $X_{ik}^{\mathcal{O}}(r) > 0$ we must have $W_{ik} = \frac{P_k}{\tau_k}$. But this would imply $\pi_{ik}W_{ik} = \frac{\pi_{ik}}{\tau_k}P_k > P_k$, which means that employing any workers in house in sector i to work on task k when needed, and leaving them idle otherwise, would be more profitable than employing the same workers to produce k for the spot market. \square

Lemma 2 instead tell us that, when input k can be outsourced at no cost, any in-house production of k is done by idle workers whose preferred task is not needed. In particular, full outsourcing is optimal when $\{\tau_k\}_{k=1}^K = 1$.

Lemma 2. *If $\tau_k = 1$, then industry i does not hire workers with k as their first task in the pecking order.*

Proof. Again by contradiction, consider any allocation ζ such that $i(\zeta) = i$ and $\mathbb{P}_{\zeta}(k) = \pi_{ik}$. If a worker \mathbf{s} allocated optimally according to ζ , then we must have

$$W_{ik}s_k \geq W_{ih}s_h \quad \forall h$$

Moreover, since $\tau_k = 1$, we must have $W_{ik} = P_k$. Therefore we also have

$$P_k s_k = W_{ik} s_k \geq \sum_{h=1}^K W_{ih} s_h \mathbb{P}_{\zeta}(h)$$

which means that employing worker \mathbf{s} to produce k for the spot market is more profitable. \square

Lemma ?? points out that outsourcing is never profitable when all workers have the same relative productivity across tasks.

Lemma 3. *Suppose the the skill density μ is non-zero only on a straight line (i.e. $\mu(\mathbf{s}) \neq 0$ only if $\frac{s_k}{s_K} = a_k \forall k \in \{1, \dots, K - 1\}$, for some given values $\{a_k\}_{k=1}^{K-1}$). Then if $\tau_k < 1$ no firm finds it profitable to buy input k on the spot market.*

Finally, Lemma 4 considers the allocation of workers based on the strength of their comparative advantage, and tells us that workers with a strong comparative advantage on some task k (as measured by their skill on k relative to the other inputs) are more likely to be outsourced. This is intuitive, as there is a high opportunity cost to having these workers employed on a task different from k .

Lemma 4. *If worker type \mathbf{s} is employed by a spot market producer of input k , then all worker types \mathbf{s}' such that $s'_k \geq s_k$ and $s'_h = s_h \forall h \neq k$ are also employed by a spot market producer of input k .*

Proof. If type \mathbf{s} is employed by a spot market producer of k , then we have

$$P_k s'_k - \max_{\zeta \in \mathcal{Z}} \sum_{h=1}^K W_{i(\zeta)h} \mathbb{P}_\zeta(h) s'_h \geq$$

$$P_k s_k - \max_{\zeta \in \mathcal{Z}} \sum_{h=1}^K W_{i(\zeta)h} \mathbb{P}_\zeta(h) s_h \geq 0$$

Therefore it is also optimal to employ \mathbf{s}' in the production of k for the spot market. \square

3.1.1 Supermodularity as a special case

So far we considered a generic skill domain \mathcal{S} , and a generic distribution $\mu \in \Delta(\mathcal{S})$. Allocation models often assume a one-dimensional skill distribution, so as to obtain a simple increasing mapping between skills and wages. In this context, a common assumption is that of supermodularity between skills and tasks (see Costinot and Vogel (2010)). That is, there is an ordering of skills and an ordering of tasks such that higher-index workers have a comparative advantage on higher-index tasks.

Let's index tasks by $k \in \{1, \dots, K\}$, and without loss of generality assume that higher-index tasks are more skill-intensive. Within our framework, a sufficient condition to obtain supermodularity is that the density μ is concentrated on a sub-space of \mathcal{S} , which can be described as a function $f : \sigma \in [0, 1] \mapsto \mathbb{R}^K$ such that, denoting the k -th component of f by f_k , we have

1. $\forall k, f_k$ is weakly increasing in σ

2. $\frac{f_k(\sigma)}{f_{k'}(\sigma)} > \frac{f_k(\sigma')}{f_{k'}(\sigma')}$ whenever $k > k'$ and $\sigma > \sigma'$

When input needs are not volatile it is straightforward to show that there is a strictly increasing allocation function M from $[0, 1]$ into $\{1, \dots, K\}$, as in Lemma 1 from Costinot and Vogel (2010). When input needs are volatile, instead, a weaker property holds, as stated in Lemma ??.

Lemma 5. *Denote by \mathcal{T} the set of allocations. There is a partial ordering \succeq on allocations $\zeta \in \mathcal{T}$ such that, if $\zeta_1 \succeq \zeta_2$, and workers σ_1 and σ_2 are allocated to ζ_1 and ζ_2 respectively, then $\sigma_1 > \sigma_2$. Moreover, the ordering \succeq becomes complete when restricted to a single sector (that is, allocations ζ_1, ζ_2 such that $i(\zeta_1) = i(\zeta_2)$ are always comparable according to \succeq). Thus positive assortative matching holds within sectors.*

Proof. See Appendix. □

3.2 Comparative statics

In this section, we study the first-order effect of changes in outsourcing costs $\{\tau_k\}$ and productivity $\{A_{ik}\}$ on marginal revenue products $\{W_{ik}\}$ and worker allocations. Given marginal revenue products and allocations, equation (12) then allows us to compute changes in the wage distribution. To set the stage for the analysis, we start by reviewing the benchmark model of biased technical change. We then extend the insights of this simpler model to our general setting with multiple inputs and final goods, and volatile input needs.

Benchmark model of biased technical change In this model there are two inputs ($i = 1, 2$) and one final good, with perfect competition in the labor and input markets. We want to study the effect of a change in the relative productivity of the two inputs on the wage of the workers who produce them. To do so, we log-linearize the market clearing conditions around the initial equilibrium.

With perfect competition in the labor market, if two workers produce the same input $i \in \{1, 2\}$ with productivities s_i and s'_i , their wages must satisfy $\frac{W(s)}{s_i} = \frac{W(s')}{s'_i} \equiv W_i$. Denoting by P_1 and P_2 the prices of the two inputs, perfect competition in the goods market implies $P_i = \frac{W_i}{A_i}$ for $i \in \{1, 2\}$.

Denote by Q_1^D and Q_2^D the quantities demanded, and denote the demand elasticity by

$$\theta \equiv -\frac{d \log \frac{Q_1^D}{Q_2^D}}{d \log \frac{P_1}{P_2}}$$

Similarly, denote by Q_1^S and Q_2^S the quantities supplied, and denote the supply elasticity by

$$\varphi \equiv \frac{d \log \frac{Q_1^S}{Q_2^S}}{d \log \frac{W_1}{W_2}}$$

If each worker can produce only one of the inputs, the supply elasticity is $\varphi = 0$. If instead all workers are equally able to produce both, then $\varphi = \infty$. In the intermediate case where workers have a comparative advantage on one of the inputs, we have $0 < \varphi < \infty$.

Finally, denote by A_1 and A_2 the productivity of labor in making the two inputs. The market clearing conditions impose that relative demand for the two inputs must equal relative supply:

$$d \log \frac{Q_1^D}{Q_2^D} = -\theta d \log \frac{P_1}{P_2} = \theta \left(d \log \frac{A_1}{A_2} - d \log \frac{W_1}{W_2} \right) \quad (13)$$

$$= d \log \frac{Q_1^S}{Q_2^S} = \varphi d \log \frac{W_1}{W_2} + d \log \frac{A_1}{A_2} \quad (14)$$

We can then solve for wages:

$$d \log \frac{W_1}{W_2} = \frac{\theta - 1}{\varphi + \theta} d \log \frac{A_1}{A_2} \quad (15)$$

Equation (15) tells us that relative wages are increasing in relative productivity if and only if the elasticity of relative demand with respect to relative productivity (θ) is greater than the elasticity of relative supply (φ), that is, if and only if inputs are substitutes. Below we will consider a more general environment, where the relative supply elasticity is $\xi_A \in [0, 1]$. It is immediate to verify that the same result holds: relative wages are increasing in relative productivity if and only if $\theta > \xi_A$.

Equation (15) also shows us that the wage response is dampened when the relative supply of workers responds to prices ($\varphi > 0$). In the extreme case where workers can freely move across inputs ($\varphi = \infty$) we have $d \log \frac{W_1}{W_2} = 0$, because input prices must remain equal.

General setting: notation To study comparative statics, we log-linearize the model around an initial equilibrium. As a simplifying assumption, we impose that the set of inputs for which each industry i has a strictly positive outsourcing share does not change in a neighborhood of the initial equilibrium.

It is convenient to redefine the set of inputs to include (a) those that are traded on the spot market, and (b) a new set of distinct inputs, corresponding to those produced only in house by some industry i . For example, suppose that there are three industries in the economy, 1, 2 and 3. Industries 1 and 2 buy a positive share of input k on the spot market, while industry 3 produces it entirely in house. In this case we include two distinct versions of input k in our set. The first has price P_k (i.e. the spot market price), and quantity equal to the total units of k used by industries 1 and 2. The second has price $W_{ik} < \tau_k P_k$,² and quantity equal to the total units of k used by industry 3.

With some abuse of notation we still index inputs by k , which now runs from 1 to $\hat{K} \geq K$. For each input we denote the total quantity produced by X_k , and the price by W_k . We stack all quantities and prices into vectors $\mathbf{X} \in \mathbb{R}^{\hat{K}}$ and $\mathbf{W} \in \mathbb{R}^{\hat{K}}$. Likewise, we denote by $\boldsymbol{\tau} \in \mathbb{R}^{\hat{K}}$ the column vector of input-specific trade costs, with the convention that $\tau_k \equiv \infty$ for inputs k that are only produced in-house. We instead maintain the distinction between productivity changes across sector-input pairs, and we denote by $\mathbf{A} \in \mathbb{R}^{N \times K, 1}$ the column vector of sector-input specific productivities.

To set the stage for the derivations below, Table 3 introduces notation for local elasticities of substitution, consumption shares, input shares in production, and input shares in total GDP.

Consumption shares	$\beta_i = \frac{P_i C_i}{\sum_j P_j C_j}$
Input shares in production	$\Omega_{ik} = \frac{W_{ik} X_{ik}}{P_i Y_i}$
Input shares in total GDP	$\Lambda_k = \frac{W_k X_k}{\sum_j P_j C_j} = \sum_i \beta_i \Omega_{ik}$
Elasticities of substitution in production	$\theta_{kh}^i = \frac{d \log \frac{X_{ik}}{X_{jh}}}{d \log \frac{G_{jh}^i}{G_{ik}^i}}$
Elasticities of substitution in consumption	$\sigma_{ij} = \frac{d \log \frac{c_i}{c_j}}{d \log \frac{c_j^i}{c_i^j}}$

Table 3: Input shares and elasticities of substitution

²Note that the product $\tau_k P_k$ could be infinity, if there is no spot market for input k .

Demand Moving from the first order conditions (6) and consumer optimization, we can derive a parallel to equation (13) in our more general model:

$$d \log \mathbf{X} - \mathbf{1} d \log \bar{X} = \Theta_A d \log \mathbf{A} - \Theta (d \log \mathbf{W} - d \log \boldsymbol{\tau}) \quad (16)$$

where

$$\Theta_{kh} \equiv \frac{\partial \log \frac{X_k}{\bar{X}}}{\partial \log W_h} = \sum_i \beta_i \omega_{ik} \left[\omega_{ih} \theta_{kh}^i (\mathbb{I}_{k=h} - 1) + \sum_j \beta_j \sigma_{ij} (\omega_{ik} - \omega_{jk}) (\omega_{ih} - \omega_{jh}) \right]$$

$$(\Theta_A)_{k,ih} \equiv \frac{\partial \log \frac{X_k}{\bar{X}}}{\partial \log A_{ih}} = \beta_i \omega_{ik} \left[\omega_{ih} \theta_{kh}^i (\mathbb{I}_{k=h} - 1) + \sum_j \beta_j \sigma_{ij} (\omega_{ik} - \omega_{jk}) (\omega_{ih} - \mathbb{I}_{i=j} \omega_{jh}) \right]$$

The first term in the expressions for Θ and Θ_A describes the producers' substitution across inputs, while the second term describes the consumers' substitution across final goods. Note that demand responds in the same way to productivity and outsourcing costs. Indeed, a fall in the outsourcing cost of input k is like a k -biased productivity change which hits all sectors that outsource a positive share of good k .

Lemma 6. *It holds that $\boldsymbol{\Lambda}^T \Theta = \boldsymbol{\Lambda}^T \Theta_A = \mathbf{0}^T$, and $\Theta \mathbf{1} = \Theta_A \mathbf{1} = \mathbf{0}$. The first set of equalities follows immediately from the fact that equation (16) solves for relative demand changes. The second set of equalities instead states that relative input demand does not change when all prices and productivities change proportionately ($d \log \mathbf{W} \propto \mathbf{1}$ and $d \log \mathbf{A} \propto \mathbf{1}$).*

Supply The total supply of input k is equal to

$$X_k = \int_{\mathcal{S}} \sum_{\zeta \in \mathcal{Z}} \xi(\mathbf{s}, \zeta) X_k(\mathbf{s} | \zeta) d\mu(\mathbf{s})$$

where we denoted by

$$X_k(\mathbf{s} | \zeta) \equiv A_{i(\zeta)k} \mathbb{P}_k(\zeta) s_k$$

The change in the supply of input k is given by

$$d \log X_k = \sum_{\zeta} \int_{\mathcal{S}} \frac{\xi(\mathbf{s}, \zeta) X_k(\mathbf{s} | \zeta)}{X_k} d\mu(\mathbf{s}) d \log A_{i(\zeta)k} + \frac{X_k^{\mathcal{O}}}{X_k} d \log \tau_k + \int_{\mathcal{S}} \sum_{\zeta \text{ s.t. } i(\zeta)=i} \frac{X_k(\mathbf{s} | \zeta)}{X_k} d\xi(\mathbf{s}, \zeta) d\mu(\mathbf{s})$$

The first two terms describe the mechanical effect of higher productivity on input supply. Total supply increases proportionately to the share of the input that has become more productive (either through biased productivity change in some sectors, $d \log A_{ik}$, or through lower outsourcing costs τ_k). The third term describes a reallocation effect, captured by changes in the probability $\xi(\mathbf{s}, \zeta)$ that each worker \mathbf{s} is employed according to the allocation ζ . From the optimal worker allocation (11), ξ depends on task wages. Moreover, whenever industry i purchases input k on the spot market, ξ also depends on outsourcing costs through the condition $W_{ik} = \frac{P_k}{\tau_k}$. Thus we have

$$d\xi(\mathbf{s}, \zeta) = \nabla_W \xi(\mathbf{s}, \zeta) d \log \mathbf{W} + \nabla_{\tau} \xi(\mathbf{s}, \zeta) d \log \tau$$

Lemma 7 follows immediately from the envelope theorem, which guarantees that changes in the allocation of workers have no effect on aggregate output around an efficient equilibrium.

Lemma 7. *Around an efficient equilibrium, it holds that*

$$\sum_{\zeta} \int_{\mathcal{S}} \sum_k \Lambda_k \frac{X_k(\mathbf{s} | \zeta)}{X_k} d\xi(\mathbf{s}, \zeta) d\mu(\mathbf{s}) = 0$$

Using Lemma 7 we can write a parallel to equation (14) in our more general setting:

$$d \log X_k - \Lambda^T d \log \mathbf{X} = \Phi d \log \mathbf{W} + [\Phi_{\tau} + \Xi_{\tau}] d \log \tau + \Xi_A d \log \mathbf{A} \quad (17)$$

where

$$\begin{aligned}
\Phi_{kh} &\equiv \sum_{\zeta} \int_{\mathcal{S}} \frac{X_k(\mathbf{s} | \zeta)}{X_k} \frac{\partial \xi(\mathbf{s}, \zeta)}{\partial \log W_h} d\mu(\mathbf{s}) \\
(\Xi_{\tau})_{kh} &\equiv \frac{X_k^{\mathcal{O}}}{X_k} \mathbb{I}_{k=h} - \sum_h \Lambda_h \frac{X_h^{\mathcal{O}}}{X_h} \\
(\Phi_{\tau})_{kh} &\equiv \sum_{\zeta} \int_{\mathcal{S}} \frac{X_k(\mathbf{s} | \zeta)}{X_k} \frac{\partial \xi(\mathbf{s}, \zeta)}{\partial \log \tau_h} d\mu(\mathbf{s}) \\
(\Xi_A)_{k,ih} &\equiv \sum_{\zeta \text{ s.t. } i(\zeta)=i} \int_{\mathcal{S}} \left[\xi(\mathbf{s}, \zeta) \left[\frac{X_k(\mathbf{s} | \zeta)}{X_k} \mathbb{I}_{k=h} - \sum_h \Lambda_h \frac{X_h(\mathbf{s} | \zeta)}{X_h} \right] \right] d\mu(\mathbf{s})
\end{aligned}$$

Lemma 7 implies that $\mathbf{\Lambda}^T \Phi = \mathbf{\Lambda}^T \Phi_{\tau} = \mathbf{0}^T$, and we have $\mathbf{\Lambda}^T \Xi_A = \mathbf{0}^T$ by construction. Moreover, it is immediate to verify that $\Phi \mathbf{1} = \Xi_A \mathbf{1} = \mathbf{0}$.

Note that changes in productivity only have a direct effect on input supply (i.e. supply increases proportionately with productivity), while changes in τ have both a direct effect (Ξ_{τ}) and an effect through worker reallocation (Φ_{τ}). For both \mathbf{A} and τ , the direct effect is proportional to the share of the input that is exposed to the productivity change. This share can be less than one, if a biased productivity change only affects a subset of the sectors, or if some units of the input are produced in house.

The reallocation effects Φ and Φ_{τ} depend on changes in the probabilities $\xi(\mathbf{s}, \zeta)$. First note that we can divide the skill set \mathcal{S} into a finite number of open subsets \mathcal{S}_{ζ} , characterized by a unique optimal allocation ζ such that $\xi(\mathbf{s}, \zeta) = 1$ for $\mathbf{s} \in \mathcal{S}_{\zeta}$. On the boundary of \mathcal{S}_{ζ} , we can have $\xi(\mathbf{s}, \zeta) \in (0, 1)$. If this boundary does not contain any mass point of μ , the integral

$$\int_{\mathcal{S}_{\zeta}} \frac{X_k(\mathbf{s} | \zeta)}{X_k} d\xi(\mathbf{s}, \zeta) d\mu(\mathbf{s})$$

can be written as a surface integral over the boundary $\delta(\mathcal{S}_{\zeta})$ of the set \mathcal{S}_{ζ} :

$$\int_{\mathcal{S}_{\zeta}} \frac{X_k(\mathbf{s} | \zeta)}{X_k} d\xi(\mathbf{s}, \zeta) d\mu(\mathbf{s}) = \int_{\delta(\mathcal{S}_{\zeta})} \frac{X_k(\mathbf{s} | \zeta)}{X_k} \mu(\mathbf{s}) \mathbf{v}(\mathbf{s}) \cdot d\mathbf{\Sigma}$$

where $\mathbf{v}(\mathbf{s})$ is the shift in the boundary of \mathcal{S}_{ζ} at point \mathbf{s} , and $d\mathbf{\Sigma}$ is the unit normal component of the surface $\delta(\mathcal{S}_{\zeta})$.

If instead the boundary contains a mass point $\hat{\mathbf{s}}$, we must include an additional term:

$$\int_{\mathcal{S}_\zeta} \frac{X_k(\mathbf{s} | \zeta)}{X_k} d\xi(\mathbf{s}, \zeta) d\mu(\mathbf{s}) = \int_{\delta(\mathcal{S}_\zeta)} \frac{X_k(\mathbf{s} | \zeta)}{X_k} \mu(\mathbf{s}) \mathbf{v}(\mathbf{s}) \cdot d\boldsymbol{\Sigma} + \frac{X_k(\hat{\mathbf{s}} | \zeta)}{X_k} d\xi(\hat{\mathbf{s}}, \zeta) \mu(\mathbf{s}) \quad (18)$$

In general, $\mathbf{v}(\mathbf{s})$ is a complicated function of task wages \mathbf{W} and trade costs $\boldsymbol{\tau}$. However, equation (18) highlights that the reallocation effects in Φ and Φ_τ are bound to be small when the mass of workers on the boundary of the sets \mathcal{S}_ζ is small.

Equilibrium Combining the demand and supply equations (16) and (17), and normalizing $P_C = 1$,³ we obtain a generalized version of the equilibrium condition (15):

$$d \log \mathbf{W} = \left(\mathbf{1} \boldsymbol{\Lambda}^T + \Phi + \Theta \right)^{-1} [(\Theta_A - \Phi_A) d \log A + (\Theta_\tau - (\Phi_\tau + \Xi_\tau)) d \log \tau] \quad (19)$$

Like in the simple benchmark model, the response of relative wages to productivity and trade costs depends on the difference between demand and supply elasticities ($\Theta_A - \Phi_A$ and $\Theta_\tau - (\Phi_\tau + \Xi_\tau)$). With respect to the benchmark two-input model, a change in the productivity of one input can have asymmetric effects on the price of the other inputs. DISCUSS.

The next two sections illustrate the general results in two special cases. These special cases introduce additional assumptions on the skill distribution, which allow us to characterize more precisely the supply elasticities Φ and Φ_τ . This in turn generates sharp predictions for the response of relative wages and spot market trade to changes in \mathbf{A} and $\boldsymbol{\tau}$.

4 Cost of outsourcing and hollowing out of the wage distribution

In this Section I consider a special case of the general framework, designed to illustrate how lower fragmentation costs can lead to a hollowing out of the wage distribution. To simplify the analysis, I restrict the economy to

³This is equivalent to imposing $\boldsymbol{\Lambda}^T \mathbf{W} \equiv 1$.

have one final good and a small number of inputs. This is without loss of generality. More importantly, I make assumptions about the skill distribution that are crucial for the results.

First, the skill density μ is one-dimensional, with supermodularity (as illustrated in Section ??). Besides allowing for a complete ordering of skills, the supermodularity assumption implies a specific pattern of comparative advantage along the skill distribution. Top-skill and bottom-skill workers have a strong comparative advantage on the hardest and easiest task respectively, whereas mid-skill workers have similar skill on a range of tasks. Second, I also assume that the mass of workers is concentrated around the low- to middle-skill range, while there are relatively few top-skill workers.

Under these assumptions, I show that a uniform fall in outsourcing costs across all inputs generates a hollowing out of the wage distribution, in the sense of (Autor and Dorn (2013)). The key intuition is that firms have a stronger incentive to outsource top- and bottom-skill workers, because they have a strong comparative advantage on one specific task. Therefore it is wasteful to employ them on other suboptimal tasks. Hence, even if the outsourcing technology improves uniformly across inputs, this only increases the productivity of top- and bottom-skill workers.

As explained in Section 3, relative wages are not necessarily increasing in relative productivity, due to a tension between a productivity effect and a supply effect. The productivity effect prevails when the demand elasticity is smaller than the initial outsourcing share. Thus, at least when initial outsourcing shares are small, lower outsourcing costs increase the relative wage of top-skill and bottom-skill workers compared to mid-skill ones. Moreover, the assumption that the skill density is concentrated towards low- to mid-skill workers implies that skill-intensive inputs have a smaller initial outsourcing share. As a result top-skill workers benefit from the same productivity effect as bottom-skill ones, but they face a smaller supply effect. This increases the skill premium.

Altogether mid-skill workers lose compared to all the others, and top-skill workers gain compared to the least skilled ones.

4.1 Environment

There is one final good in the economy ($N = 1$), produced with three intermediate inputs ($K = 3$). The production function for final goods is CES with elasticity of substitution θ :

$$Y = \left(X_1^{\frac{\theta-1}{\theta}} + X_2^{\frac{\theta-1}{\theta}} + X_3^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}$$

Agents are characterized by a vector of skills $\mathbf{s} = (s_0, s_1, s_2)$. All agents have the same productivity $s_0 = 1$ on task 0 (the low-skill task), while they differ in their productivity on inputs 1 and 2, which varies in $[0, \infty)$.

The skill density μ is concentrated on a one-dimensional subset of the (s_1, s_2) space, which can be described as an increasing and convex function $s_1 = f(s_2)$, implying a comparative advantage for high-skill workers on task 2. Along this line, the density is a continuous function $g(s_2)$.

All inputs are needed with the same probability π , and they face the same iceberg cost τ . Throughout this section we focus on an equilibrium where only inputs 0 and 2 are sold on the spot market, and we denote by P_0 and P_2 their spot market prices. Input 0 is the numeraire, with $P_0 \equiv 1$. The task wage W_1 is the marginal revenue product of input 1 for the final good firm. The marginal revenue products of inputs 0 and 2 are $\frac{1}{\tau}$ and $\frac{P_2}{\tau}$ respectively, following equation (7).

Below I derive analytically the first-order change in relative wages in response to changes in the outsourcing cost τ , and show how it depends on the density g . To go beyond the first-order approximation I also provide a quantitative illustration.

4.2 Optimal allocation of workers

Figure ?? represents the equilibrium allocation of workers to tasks, for different values of the outsourcing cost τ .⁴

The Appendix reports the equations that define the various regions.

⁴To construct Figure ?? we compute equilibrium prices given τ and the skill density μ defined above. Given prices, however, we can compute the optimal allocation over the entire skill space, as represented in the figure.

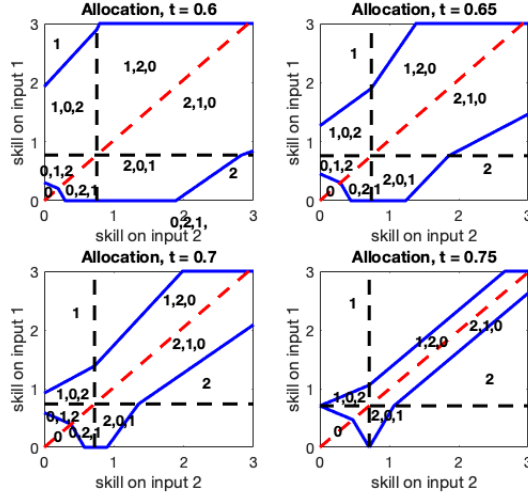


Figure 1: Optimal allocation of workers

As the figure illustrates, firms always keep in house the workers who have similar skills on all tasks (i.e. around the point $s_1 = s_2 = 1$). For low-enough outsourcing costs, they also keep in house workers with similar skill on tasks 1 and 2 (i.e. those around the diagonal $s_1 = s_2$), or on task 0 and one of the other tasks (i.e. around the lines $s_1 = 1$ or $s_2 = 1$). Each worker is employed on the task where he has comparative advantage among those needed, and remains idle with probability $1 - (1 - \pi)^2$. The workers who have strong comparative advantage on a specific task instead are employed by spot market producers. These are the workers with high skill on task 1 or task 2, or very low skill on both.

Figure 2 also depicts the line along which the density μ is concentrated. With supermodularity the workers with strong comparative advantage are those at the extremes of the distribution, who therefore are outsourced. The workers in the middle of the distribution are closer to the point $s_1 = s_2 = 1$, so they have more similar skill in all tasks, and for low-enough outsourcing costs they are kept in house.

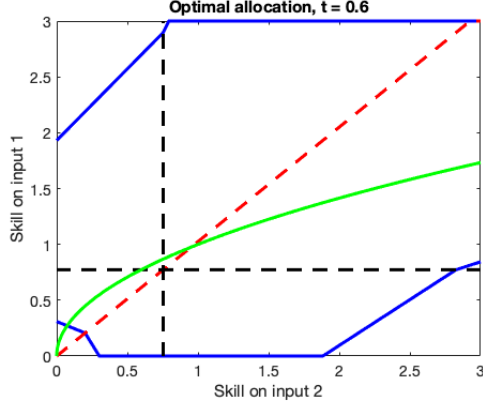


Figure 2: Optimal allocation and skill distribution

With a one-dimensional density, the agents' optimal allocation can be expressed as a function of their skill on input 2, s_2 . From now on we restrict attention to an equilibrium where only inputs 0 and 2 are sold on the spot market, while input 1 is entirely produced in house, and perform comparative statics around this equilibrium. In this case the allocation is characterized by five cutoff points for s_2 , which we label s_0^* , s_1^* , s_2^* , s_3^* , and s_4^* , as follows:

$$\left\{ \begin{array}{ll} s_2 < s_0^* & \text{s outsourced on 0} \\ s_0^* < s_2 < s_1^* & \text{s in house on 0, 1, 2} \\ s_1^* < s_2 < s_2^* & \text{s in house on 1, 0, 2} \\ s_2^* < s_2 < s_3^* & \text{s in house on 1, 2, 0} \\ s_3^* < s_2 < s_4^* & \text{s in house on 2, 1, 0} \\ s_2 > s_4^* & \text{s outsourced on 2} \end{array} \right.$$

The Appendix reports explicit expressions for the thresholds.

4.3 Comparative statics: outsourcing cost

We now consider how reducing spot market trade costs (i.e. increasing τ) affects the wage distribution. The response of relative input prices (W_1 and P_2) is key to understanding relative wages. I first show that, when initial

outsourcing shares are small, the relative price of the mid-skill input 1 increases relative to the low-skill input 0, but less than one-for-one with τ . As a consequence, the wage of mid-skill workers falls relative to the least skilled workers, who are outsourced on input 0. I then show that, if the initial outsourcing share of input 2 is smaller than that of input 0, the price of the high-skill input 2 also increases relative to input 0. Thus, the wage of top-earners relative to bottom-earners also increases.

The response of relative prices depends on the difference between demand and supply elasticities, as described by the equilibrium equation (19). With respect to the general derivation in Section 3, it is convenient to normalize prices and quantities relative to input 0, because ultimately we will solve for the change in wages relative to the least skilled workers. Accordingly, we adopt a small change in notation: while in Section 3 we denoted by Φ the supply elasticities relative to the aggregate, here we denote by \mathcal{E} the supply elasticities with respect to input 0. The two are related as follows:

$$\mathcal{E} = \Phi_{1:2,1:2} - \Phi_{1:2,0}\mathbf{1}^T$$

With this notation, price changes relative to input 0 are given by

$$\frac{d \log W_1}{d \log \tau} = \frac{\left(\frac{X_0^O}{X_0} - \mathcal{E}_1^\tau\right) - \frac{\mathcal{E}_{12}^p}{\mathcal{E}_{22}^p + \theta} \left(\frac{X_0^O}{X_0} - \frac{X_2^O}{X_2} - \mathcal{E}_2^\tau\right)}{\mathcal{E}_{11}^p + \theta - \frac{\mathcal{E}_{12}^p}{\mathcal{E}_{22}^p + \theta} \mathcal{E}_{21}^p} \quad (20)$$

$$\frac{d \log P_2}{d \log \tau} = \frac{\left(\frac{X_0^O}{X_0} - \frac{X_2^O}{X_2} - \mathcal{E}_2^\tau\right) - \frac{\mathcal{E}_{21}^p}{\mathcal{E}_{11}^p + \theta} \left(\frac{X_0^O}{X_0} - \mathcal{E}_1^\tau\right)}{\mathcal{E}_{22}^p + \theta - \frac{\mathcal{E}_{21}^p}{\mathcal{E}_{11}^p + \theta} \mathcal{E}_{12}^p} \quad (21)$$

The reallocation effects \mathcal{E}^p and \mathcal{E}^τ are complicated functions of W_1 , P_2 and τ . I report them in the Appendix. A sufficient condition for $\frac{d \log X}{d \log \tau} > 0$ for $X \in \{W_1, P_2\}$ is

$$\frac{X_0^O}{X_0} - \frac{X_i^O}{X_i} - \mathcal{E}_i^\tau > 0 \text{ and } \frac{\frac{X_0^O}{X_0} - \frac{X_i^O}{X_i} - \mathcal{E}_i^\tau}{\frac{X_0^O}{X_0} - \frac{X_j^O}{X_j} - \mathcal{E}_j^\tau} > \frac{\mathcal{E}_{ij}^p}{\mathcal{E}_{jj}^p + \theta} \quad (22)$$

Condition (22) requires that the outsourcing share on input 0 is large enough, so that a decline in outsourcing costs increases the supply of input 0 relative to the others. This is more likely to hold when the density g is concentrated around the lower cutoffs s_0^* and s_1^* , because in this case the outsourcing share of input 2 is smaller than that of input 0 ($0 \leq \frac{X_2^O}{X_2} \leq \frac{X_0^O}{X_0}$). In addition, condition (22) also requires that cross-price effects on input i arising from

changes in the relative supply of input j are small relative to the direct effect of τ on i . This is the case when \mathcal{E}_{ij}^p is small compared to \mathcal{E}_{jj}^p .

In turn, whenever $\frac{d \log P_2}{d \log \tau} > 0$ a decline in outsourcing costs raises the wage of the workers outsourced on input 2 increases relative to those outsourced on input 0. In fact, the relative productivity of these workers is constant in τ , while the relative price of input 2 increases. Formally, equation (12) implies that the wage of the highest earner $\bar{\mathbf{s}}$ (who produces input 2 for the spot market) relative to the lowest earner $\underline{\mathbf{s}}$ (who produces input 0 for the spot market) changes proportionately to the price of input 2:

$$d \log \frac{w(\bar{\mathbf{s}})}{w(\underline{\mathbf{s}})} = d \log P_2 > 0$$

The relative productivity of in-house vs outsourced workers instead is not constant in τ . Thus, a decline in outsourcing costs decreases the wage of the least skilled in-house workers relative to outsourced workers on input 0 whenever the price of input 1 increases proportionately less than the productivity of outsourced workers ($\frac{d \log W_1}{d \log \tau} > 1$). This is the case if and only if

$$\theta + \mathcal{E}_{11}^p > \left(\frac{X_0^O}{X_0} - \mathcal{E}_1^\tau \right) - \frac{\mathcal{E}_{12}^p}{\mathcal{E}_{22}^p + \theta} \left(\frac{X_0^O}{X_0} - \frac{X_2^O}{X_2} - \mathcal{E}_2^\tau - \mathcal{E}_{21}^p \right) \quad (23)$$

Note that, different from the traditional model of biased technical change, this condition can be satisfied even when inputs are complements ($\theta < 1$). To the extent that reallocation effects are small, condition (23) is satisfied when the demand elasticity θ is larger than the initial outsourcing share of input 0. This is because changes in τ only affect outsourced units, therefore they have a smaller supply effect.

Again following Equation (12), the wage of in-house workers relative to the lowest earner $\underline{\mathbf{s}}$ is given by

$$d \log \frac{w(\mathbf{s})}{w(\underline{\mathbf{s}})} = s_2 \frac{\tau P_2 \mathbb{P}(\mathbf{s} \text{ on } 2)}{w(\mathbf{s})} d \log P_2 + s_1 \frac{P_1 \mathbb{P}(\mathbf{s} \text{ on } 1)}{w(\mathbf{s})} (d \log W_1 - 1) - \frac{\tau \mathbb{P}(\mathbf{s} \text{ on } 0)}{w(\mathbf{s})}$$

When $d \log W_1 < 1$, this expression is negative for the least-skilled in-house workers (with small s_2). Thus, these workers lose compared to those at the top and bottom end of the distribution.

In a numerical example, Figure ?? illustrates the evolution of the wage distribution relative to the least-skilled

worker as τ increases from 0.6 to 0.9. The rest of the calibration is summarized in Table ??.

Volatility of needs	$\pi = 0.5$
Elasticity of substitution	$\theta = 0.1$
skill density	$g \sim \mathcal{N}(0.7, 0.5)$, truncated at zero
supermodularity	$f = \sqrt{\cdot}$

Table 4: Model calibration

As outsourcing becomes cheaper, the least skilled workers that were previously in house are outsourced on input 0. Since input 0 has become cheaper relative to the others, these workers face a wage cut. By contrast, the most skilled workers that were previously in house are outsourced on input 2. The relative price of input 2 increases, therefore these workers gain. In-house workers are in between. By continuity wages fall for those at the bottom of the distribution, whereas they increase for those at the top.

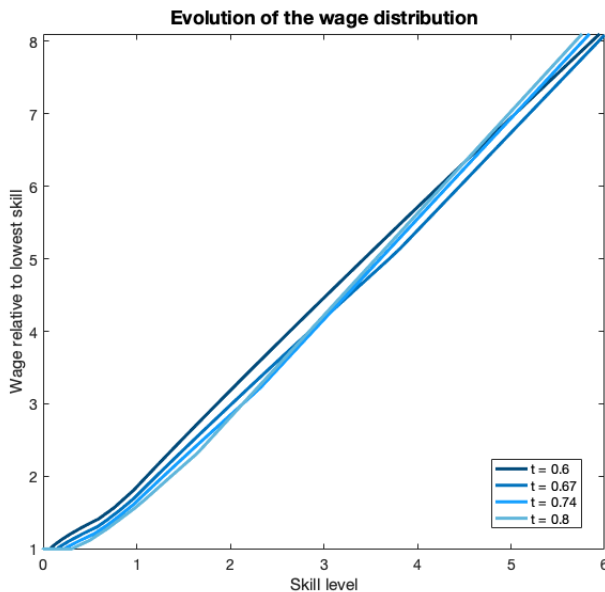


Figure 3: Evolution of the wage distribution after a fall in outsourcing costs

5 Skill-biased technical change and outsourcing

In this section I specialize the general model to study how skill-biased technical change affects the incentives to engage in outsourcing. Essentially, a skill-biased productivity change increases the skill premium, thereby making it more costly for firms to employ high-skill workers on tasks where they do not have comparative advantage. This automatically creates an incentive to split the production of high and low skill tasks between different industries.

To highlight the connection with low-skill outsourcing, I embrace a specific interpretation of the suboptimal tasks that high-skill workers might be required to perform when hired in-house. That is, top-skill workers might occasionally have to manage low-skill ones, dealing with hiring/firing, assigning tasks, workplace issues, etc. To the extent that hiring and workplace issues happen infrequently, it might not be worth it for firms to have dedicated human resource (HR) managers in-house. For example, an economics department likely won't need to solve issues with its cleaning crew very often, therefore it prefers not to hire an HR person specifically for this task. The department then has two alternatives: asking a faculty member to manage the cleaning crew when needed, or hiring a cleaning company.

Note that whether the cleaning crew is hired directly by the university or through a cleaning company has no implication for where the cleaners work. It only allows multiple departments (or more broadly multiple final good producers) to share the HR people who supervise and manage their low-skill workers. At the same time this kind of outsourcing is more prevalent when the skill premium is higher, because the opportunity cost of employing high-paid workers on less valuable HR tasks is higher. According to this story, outsourcing is caused by the increase in the skill premium rather than being the cause of it, as suggested in previous work.

The results in this section raise a word of caution against imposing a specific direction of causality on the empirical correlation between outsourcing and wage differentials. It is outside the scope of this paper to evaluate the relative importance of reallocation frictions, as in my model, versus bargaining frictions and firm premia, as for example in Goldschmidt and Schmieder (2017). However, at the end of this section I show that skill-biased technical change has quantitatively realistic implications for the joint evolution of the outsourcing share, the skill premium, and the expenditure share on manufacturing.

5.1 Environment

There are three final goods (manufacturing, tech, and administrative services) and three inputs in the economy (a low-skill bundle, HR services, and engineering services). The low-skill bundle further combines low-skill labor with HR services.

Final consumers have CES preferences over manufacturing, administrative services and tech products, given by

$$C = \left(\beta_T^{\frac{1}{\theta}} C_T^{\frac{\theta-1}{\theta}} + \beta_M^{\frac{1}{\theta}} C_M^{\frac{\theta-1}{\theta}} + \beta_A^{\frac{1}{\theta}} C_A^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}$$

Manufacturing firms use as inputs engineering services, $X_{M,E}$, and a low-skill bundle, Y_{LS} :

$$Y_M = \left(\omega_{LS}^{\frac{1}{\sigma}} Y_{LS}^{\frac{\sigma-1}{\sigma}} + \omega_E^{\frac{1}{\sigma}} X_{M,E}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

In turn, the low-skill bundle combines the low-skill input X_{LS} with the HR input X_{HR} :

$$Y_{LS} = \left(\alpha_{LS}^{\frac{1}{\epsilon}} X_{LS}^{\frac{\epsilon-1}{\epsilon}} + \alpha_{HR}^{\frac{1}{\epsilon}} X_{HR}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}$$

Tech firms only use engineering services, and they have productivity A_T :

$$Y_T = A_T X_{T,E}$$

Finally, administrative services are produced one for one with the HR input:

$$Y_A = X_{A,HR}$$

Workers are endowed with a vector of skills, whose components correspond to their productivity on low-skill labor, HR services, and engineering services. The skill distribution is concentrated on three mass points: $\mathbf{s} = (1, 0, 0)$, with mass $\bar{L}S$, $\mathbf{s} = (1, 1, 0)$, with mass $\bar{H}R$, and $\mathbf{s} = (1, 1, 1)$, with mass \bar{E} . This means that there are three types of workers in the economy. Low-skill workers can only provide low-skill labor, managers can provide both low-skill

labor and HR services, and engineers can provide all three inputs. Conditional on being able to provide an input, all workers have the same productivity.

All inputs in all sectors are needed all the time and face no iceberg trade cost, except for HR workers in the low-skill bundle, who are needed with probability π and face iceberg cost τ . We further assume that low-skill workers are employed by the same firm as their managers.⁵ This way low-skill workers are outsourced whenever their manager is outsourced, even though they ultimately serve only one manufacturing firm. The iceberg cost is incurred whenever the same manager serves multiple manufacturing firms.

5.2 Optimal allocation of workers

In this economy there can be either three or four relevant input prices, depending on whether manufacturing firms purchase some HR services on the spot market. If they do, there are only three relevant prices: P_E (the price of engineering services, which by condition (7) equals its marginal revenue product in both tech and manufacturing); P_{LS} (the price of low-skill labor, which is equal to the marginal revenue product of low-skill workers in the low-skill bundle); and P_{HR} (the price of administrative services). By condition (7) the marginal revenue product $W_{M,HR}$ of HR services in the low-skill bundle is $W_{M,HR} = \frac{P_{HR}}{\tau}$. If instead manufacturing firms produce HR services entirely in-house, there is a fourth relevant price $W_{M,HR} < \frac{P_{HR}}{\tau}$.

Given that workers with higher skill are able to produce less skill-intensive inputs, but not vice versa, prices must be increasing in skill intensity. Moreover we must always have $W_{M,HR} \geq P_{HR}$ in equilibrium, otherwise all HR workers and engineers would want to work for an HR service provider.⁶ Therefore we must have $P_{LS} \leq P_{HR} \leq W_{M,HR} \leq P_E$. All throughout, we normalize $P_{LS} = 1$.

Figure 4 depicts the optimal allocation of workers to tasks given relative prices.

⁵Manufacturing firms face the same cost when they hire low-skill workers directly or when they purchase their services from an external company, because we assumed $\pi_{LS} = \tau_{LS} = 1$.

⁶It is straightforward to see that, if $P_{M,HR} < P_{HR}$, then we also have $P_{M,HR} < \frac{P_{HR}}{\tau}$, so that manufacturing firms must produce HR services in-house. However, no HR worker would want to be employed in manufacturing and earn $\pi P_{M,HR} + (1 - \pi) P_{LS} < P_{HR}$. Likewise, it is optimal for manufacturing firms to employ engineers on HR only if $P_E = P_{M,HR}$. Therefore we would also have $P_E < P_{HR}$, but then all engineers would want to work in for an HR supplier.

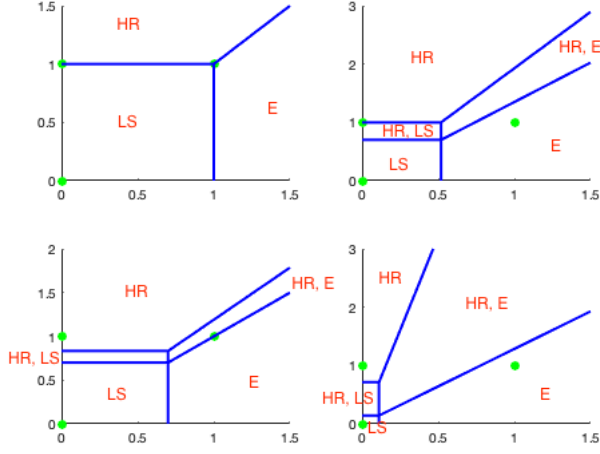


Figure 4: Optimal allocation of workers to tasks

The possible scenarios are as follows:

$$\left\{ \begin{array}{ll}
 1 = P_{HR} = W_{M,HR} = P_E, \tau = 1 & M \text{ employs workers indifferently on any task they can produce (top left panel)} \\
 1 < W_{M,HR} < \frac{P_{HR}}{\tau} \text{ and } P_E & M \text{ hires } HR \text{ workers in house, and employs them on } LS \text{ when idle (top right panel)} \\
 1 < W_{M,HR} = \frac{P_{HR}}{\tau} < P_E & M \text{ outsources } HR \text{ or produces it with in-house } HR \text{ (top right panel)} \\
 1 < W_{M,HR} = P_E < \frac{P_{HR}}{\tau} & M \text{ employs } E \text{ on } HR \text{ when needed (bottom left panel)} \\
 1 < W_{M,HR} = \frac{P_{HR}}{\tau} = P_E & M \text{ outsources } HR \text{ or produces it with in-house } E \text{ (bottom left panel)} \\
 1 < W_{M,HR} = \frac{P_{HR}}{\tau} < P_E & M \text{ fully outsources } HR \text{ (bottom right panel)}
 \end{array} \right.$$

If HR workers and engineers are scarce enough (so that $P_E > 1$ and $W_{M,HR} > 1$), manufacturing firms do not find it profitable to hire HR workers in house. Rather, they either employ engineers on HR tasks when needed, or they outsource the HR tasks. We assume that this is the case, and show that outsourcing in the manufacturing sector is increasing in the productivity of the tech sector.

5.3 Comparative statics: tech productivity

We start from an initial equilibrium with $W_{M,HR} = P_E < \frac{P_{HR}}{\tau}$, where manufacturing firms do not engage in outsourcing, and study the effect of a change in the tech productivity A_T . Denote by $\eta_{HR} = \frac{X_{M,HR}^O}{X_{M,HR}}$ the outsourcing share in manufacturing.⁷ We will show that the economy goes through three phases, represented in the two bottom subplots of figure 4:

$$\begin{cases} P_E = W_{M,HR} < \frac{P_{HR}}{\tau}, \eta_{HR} = 0 & \text{phase 1} \\ P_E = W_{M,HR} = \frac{P_{HR}}{\tau}, \eta_{HR} \in (0, 1) & \text{phase 2} \\ W_{M,HR} = \frac{P_{HR}}{\tau} < P_E, \eta_{HR} = 1 & \text{phase 3} \end{cases}$$

To see this, we move from the market clearing conditions for engineers and HR workers:

$$\begin{cases} \left[\frac{X_{M,E}}{X_{M,LS}} + \pi(1 - \eta_{HR}) \frac{X_{M,HR}}{X_{M,LS}} + \frac{Y_T}{A_T X_{M,LS}} \right] \frac{X_{M,LS}}{LS} = \frac{\bar{E}}{LS} \\ \left[\frac{Y_{HR}}{X_{M,LS}} + \pi \eta_{HR} \frac{X_{M,HR}}{X_{M,LS}} \right] \frac{X_{M,LS}}{LS} = \frac{\bar{H}R}{LS} \end{cases} \quad (24)$$

Note that, by assumption, our initial equilibrium is in phase 1. By differentiating equation (24) it is immediate to see that, in each of the three phases, P_E is increasing in A_T , and in case 2 the outsourcing share η_{HR} is increasing in A_T .⁸ Therefore, as A_T increases, the economy travels from the first to the third phase, and the manufacturing sector goes from no outsourcing to full outsourcing of the low-skill bundle.

Figure ?? plots the evolution of the outsourcing share in manufacturing, the wage of engineers relative to HR workers, and the overall sales and employment shares of manufacturing, as a function of the relative productivity of the tech sector. The figure is based on a calibration of the model summarized in table ?. We let the tech productivity vary from 0.8 to 2.2.

⁷Denote by $\zeta_{HR} \equiv (M; HR, E)$ the allocation where engineers are hired in-house by a manufacturing firm and occasionally employed on the HR task. The outsourcing share is related with the fraction $\xi(E, \zeta_{HR})$ of engineers that are employed according to ζ_{HR} , as follows:

$$\xi(E, \zeta_{HR}) = (1 - \eta_{HR}) \frac{X_{M,HR}}{\bar{E}}$$

⁸See the Appendix for the full derivation.

HR outsourcing cost	$\tau = 0.6$
Probability to need HR	$\pi = 0.5$
Elasticity of substitution in consumption	$\theta = 3$
Preference shares in consumption	$\beta_T = 0.5, \beta_A = 0.26, \beta_M = 0.24$
Elasticity of substitution in manufacturing	$\sigma = 0.5$
Input shares in manufacturing	$\omega_E = 0.4, \omega_{LS} = 0.6$
Elasticity of substitution in low-skill bundle	$\epsilon = 0.1$
Input shares in low-skill bundle	$\alpha_{LS} = \alpha_{HR} = 0.5$
Relative mass of workers	$\mu(LS) = \frac{1}{3}, \mu(HR) = \frac{1}{3}, \mu(E) = \frac{1}{3}$

Table 5: Model calibration

In an initial range of low tech productivity, manufacturing firms do not engage in outsourcing. Instead they hire engineers and low-skill workers, and let engineers perform HR tasks when needed. As tech productivity increases, P_E also increases, up to a point where manufacturing firms find it profitable to outsource some HR tasks. During the intermediate phase when the outsourcing share is $\eta_{HR} \in (0, 1)$, HR workers fully share in the rents coming from higher tech productivity, so that the relative wage of HR workers and engineers remains constant. Finally, when tech productivity is so high that manufacturing firms find it optimal to outsource HR production entirely, the relative wage of engineers starts to increase again. All throughout, given that final goods are substitutes, higher productivity in the tech sector also raises its sales and employment shares, together with the relative wage of engineers versus low-skill workers ($\frac{W_E}{W_{LS}}$).

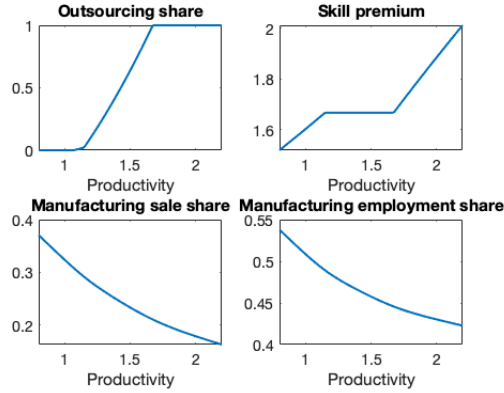


Figure 5: Effects of an increase in productivity for the tech sector

6 Conclusion

I model an economy where firms have volatile input needs, and it is costly to reallocate resources across firms. I show that in this economy there is a two-way relationship between the wage distribution and the division of production across industries.

On the one hand, skill-biased technical change increases the skill premium and creates an incentive to fragment the production of high- and low-skill tasks across distinct industries. On the other hand, a decline in the cost of outsourcing favors the workers at the top and bottom ends of the wage distribution relative to those in the middle, so that lower outsourcing costs simultaneously imply more fragmentation and a hollowing out of the wage distribution.

In order to make the general equilibrium analysis tractable I imposed assumptions which eliminate economies of scale and scope, and guarantee that external input suppliers can fully insure final good producers against input need volatility. Relaxing these assumptions would deliver a richer partial equilibrium model of fragmentation, with testable implications for the correlation between firm characteristics such as scale and scope and outsourcing choices. This would be particularly useful to develop empirical tests of the tradeoff between need volatility and trade costs, around which my model is built.

References

- ACEMOGLU, D. AND D. AUTOR (2011): “Skills, Tasks and Technologies: Implications for Employment and Earnings,” *Handbook of Labor Economics*, Elsevier.
- ACEMOGLU, D. AND P. RESTREPO (2018): “The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment,” *American Economic Review*, 108, 1488–1542.
- AUTOR, D. H. AND D. DORN (2013): “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market,” *American Economic Review*, 103, 1553–1597.
- BILAL, A. AND H. LHUILLIER (2020): “Outsourcing, Inequality and Aggregate Output,” *SSRN Electronic Journal*.

- COSTINOT, A. AND J. VOGEL (2010): “Matching and Inequality in the World Economy,” *Journal of Political Economy*, 118, 747–786, publisher: University of Chicago Press.
- COSTINOT, A., J. VOGEL, AND S. WANG (2013): “An Elementary Theory of Global Supply Chains,” *The Review of Economic Studies*, 80, 109–144, publisher: Oxford Academic.
- ECKERT, F., S. GANAPATI, AND C. WALSH (2022): “Urban-Biased Growth: A Macroeconomic Analysis,” .
- FEENSTRA, R. C. AND G. H. HANSON (1996): “Globalization, Outsourcing, and Wage Inequality,” *The American Economic Review*, 86, 240–245, publisher: American Economic Association.
- GARICANO, L. AND E. ROSSI-HANSBERG (2006): “Organization and Inequality in a Knowledge Economy,” *The Quarterly Journal of Economics*, 121, 1383–1435, publisher: Oxford University Press.
- GOLDIN, C. AND L. F. KATZ (2010): *The Race between Education and Technology*, Harvard University Press.
- GOLDSCHMIDT, D. AND J. F. SCHMIEDER (2017): “The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure*,” *The Quarterly Journal of Economics*, 132, 1165–1217.
- GROSSMAN, G. M. AND E. HELPMAN (2018): “Growth, Trade, and Inequality,” *Econometrica*, 86, 37–83, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14518>.
- GROSSMAN, G. M., E. HELPMAN, AND P. KIRCHER (2017): “Matching, Sorting, and the Distributional Effects of International Trade,” *Journal of Political Economy*, 125, 224–264.
- GROSSMAN, G. M. AND E. ROSSI-HANSBERG (2008): “Trading Tasks: A Simple Theory of Offshoring,” *American Economic Review*, 98, 1978–1997.
- HART, O. AND J. MOORE (1990): “Property Rights and the Nature of the Firm,” *Journal of Political Economy*, 98, 1119–1158, publisher: University of Chicago Press.
- KATZ, L. F. AND D. H. AUTOR (1999): “Changes in the Wage Structure and Earnings Inequality,” in *Handbook of Labor Economics*, Elsevier, vol. 3, 1463–1555.
- KATZ, L. F. AND K. M. MURPHY (1992): “Changes in Relative Wages, 1963–1987: Supply and Demand Factors,” *The Quarterly Journal of Economics*, 107, 35–78.

KIKUCHI, T., K. NISHIMURA, AND J. STACHURSKI (2018): “Span of control, transaction costs, and the structure of production chains,” *Theoretical Economics*, 13, 729–760, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/TE2376>.

KLEINMAN, B. (2023): “Wage INequality and the Spatial Expansion of Firms,” *Working Paper*.

ROSEN, S. (1981): “The Economics of Superstars,” *The American Economic Review*, 71, 845–858, publisher: American Economic Association.

SONG, J., D. J. PRICE, F. GUVENEN, N. BLOOM, AND T. VON WACHTER (2019): “Firming Up Inequality*,” *The Quarterly Journal of Economics*, 134, 1–50.

WERNERFELT, B. (2015): “The Comparative Advantages of Firms, Markets and Contracts: a Unified Theory,” *Economica*, 82, 350–367, publisher: London School of Economics and Political Science.

Appendix

Proof of Lemma 5

- Step 1: prove that the ordering is complete on \mathcal{T} . That is, for given mean, the allocations that assign higher probabilities to tasks at the extremes rather than in the middle are dominated and do not belong to \mathcal{T} . So, for each possible mean value in $[1, K]$, there is at most one allocation in \mathcal{T} with that mean value (detail: might have same allocation in different sectors, but we basically consider those the same). In other words, pecking orders in equilibrium allocations are such that as you go towards suboptimal tasks you never add an element that was in the range spanned by previous elements.
- Step 2: define a modified version \hat{f} of the productivity function f , as follows:

$$\hat{f} : \sigma \in [0, 1] \mapsto t \in \mathcal{T}$$

$$\hat{f}_t(\sigma) = \sum_k \mathbb{P}_t(k) W_{i(t)k} f_k(\sigma)$$

and show that \hat{f} is log-supermodular on \mathcal{T} . This concludes the proof.

$$\frac{\hat{f}_t(\sigma)}{\hat{f}_{t'}(\sigma)} = \frac{\sum_k \mathbb{P}_t(k) W_{i(t)k} f_k(\sigma)}{\sum_k \mathbb{P}_{t'}(k) W_{i(t')k} f_k(\sigma)} \leq \frac{\sum_k \mathbb{P}_t(k) W_{i(t)k} f_k(\sigma')}{\sum_k \mathbb{P}_{t'}(k) W_{i(t')k} f_k(\sigma')} = \frac{\hat{f}_t(\sigma')}{\hat{f}_{t'}(\sigma')}$$

$$\sum_k \frac{\mathbb{P}_t(k) W_{i(t)k} f_k(\sigma)}{\sum_k \mathbb{P}_t(k) W_{i(t)k} f_k(\sigma)} [\log(\mathbb{P}_t(k) W_{i(t)k}) + \log f_k(\sigma)] - \sum_k \frac{\mathbb{P}_{t'}(k) W_{i(t')k} f_k(\sigma)}{\sum_k \mathbb{P}_{t'}(k) W_{i(t')k} f_k(\sigma)} [\log(\mathbb{P}_{t'}(k) W_{i(t')k}) + \log f_k(\sigma)]$$

Supply elasticities in Section 4

This Section provides explicit expressions for the supply elasticities \mathcal{E}^P and \mathcal{E}^τ in Section 4. The model has a one-dimensional skill distribution, therefore supply elasticities are determined by changes in the cutoffs s_0^* , s_1^* , s_2^* , s_3^* and s_4^* . These thresholds are characterized by the following equations

$$\begin{aligned} \tau - \pi &= \pi(1 - \pi) [\tau W_1 f(s_0^*) + (1 - \pi) P_2 s_0^*] \\ 1 &= \tau W_1 f(s_1^*) \\ 1 &= P_2 s_2^* \\ 0 &= \tau W_1 f(s_3^*) - P_2 s_3^* \\ \pi(1 - \pi)^2 &= (\tau - \pi) P_2 s_4^* - \pi(1 - \pi) \tau W_1 f(s_4^*) \end{aligned}$$

Changes in the cutoffs are obtained by differentiating the equations above:

$$\begin{cases} ds_0^* &= -\frac{\tau W_1 f(s_0^*) d \log W_1 + (1 - \pi) P_2 s_0^* d \log P_2}{\tau W_1 f'(s_0^*) + (1 - \pi) P_2} + \frac{1 - \pi(1 - \pi) \tau W_1 f(s_0^*)}{\pi(1 - \pi) [\tau W_1 f'(s_0^*) + (1 - \pi) P_2]} d \log \tau \\ ds_1^* &= -\frac{f(s_1^*)}{f'(s_1^*)} [d \log W_1 + d \log \tau] \\ ds_2^* &= -s_2^* d \log P_2 \\ ds_3^* &= \frac{\tau W_1 f(s_3^*)}{P_2 - \tau W_1 f'(s_3^*)} [d \log W_1 + d \log \tau - d \log P_2] \\ ds_4^* &= \frac{\pi(1 - \pi) \tau W_1 f(s_4^*) W_1 - (\tau - \pi) P_2 s_4^* d \log P_2 - \pi[(1 - \pi)^2 + P_2 s_4^*] d \log \tau}{(\tau - \pi) P_2 - \pi(1 - \pi) \tau W_1 f'(s_4^*)} \end{cases}$$

In turn, changes in the quantity of input $i = 1, 2$ relative to input 0 are a weighted average of changes in the cutoffs,

with weights given by the relative share of input i versus input 0 produced around each cutoff:

$$\begin{aligned}\mathcal{E}_{ij}^p &\equiv \frac{\partial \log \frac{X_i}{X_0}}{\partial \log W_j} = \sum_{k=0}^4 \left[(\Delta \mathbb{P}(X_i | s_i | s_k^*) \frac{X_0}{X_i} - \Delta \mathbb{P}(X_0 | s_k^*)) \frac{g(s_k^*)}{X_0} \frac{\partial s_k^*}{\partial \log W_j} \right] \\ \mathcal{E}_i^\tau &\equiv \frac{\partial \log \frac{X_i}{X_0}}{\partial \log \tau} = \sum_{k=0}^4 \left[(\Delta \mathbb{P}(X_i | s_i | s_k^*) \frac{X_0}{X_i} - \Delta \mathbb{P}(X_0 | s_k^*)) \frac{g(s_k^*)}{X_0} \frac{\partial s_k^*}{\partial \log \tau} \right]\end{aligned}$$

where we denoted by

$$\Delta \mathbb{P}(X_i | s_k^*) \equiv \lim_{s_2 \rightarrow s_k^{*-}} \mathbb{P}(X_i | s_2) - \lim_{s_2 \rightarrow s_k^{*+}} \mathbb{P}(X_i | s_2)$$

Now we can compare \mathcal{E}_{11}^p against \mathcal{E}_{21}^p and \mathcal{E}_{22}^p against \mathcal{E}_{12}^p . Specifically, we show that the own elasticities are always larger than the cross-elasticities. The intuition is as follows: a change in the own price triggers larger changes in the cutoffs around which larger amounts of the own quantity are produced. Formally, we have

$$\begin{aligned}\mathcal{E}_{11}^p &= -(1-\pi) \left[\pi f(s_0^*) (\tau W_1)^\theta + 1 \right] \frac{g(s_0^*)}{X_0} \frac{\partial s_0^*}{\partial \log W_1} + \\ &\quad - \pi^2 \left[f(s_1^*) (\tau W_1)^\theta + 1 \right] \frac{g(s_1^*)}{X_0} \frac{\partial s_1^*}{\partial \log W_1} + \\ &\quad \pi^2 f(s_3^*) (\tau W_1)^\theta \frac{g(s_3^*)}{X_0} \frac{\partial s_3^*}{\partial \log W_1} + \\ &\quad + \pi(1-\pi) \left[f(s_4^*) (\tau W_1)^\theta - (1-\pi) \right] \frac{g(s_4^*)}{X_0} \frac{\partial s_4^*}{\partial \log W_1} \\ \mathcal{E}_{21}^p &= -(1-\pi) \left[\pi(1-\pi) s_0^* P_2^\theta + 1 \right] \frac{g(s_0^*)}{X_0} \frac{\partial s_0^*}{\partial \log W_1} + \\ &\quad - \pi^2 \frac{g(s_1^*)}{X_0} \frac{\partial s_1^*}{\partial \log W_1} + \\ &\quad - \pi^2 s_3^* P_2^\theta \frac{g(s_3^*)}{X_0} \frac{\partial s_3^*}{\partial \log W_1} + \\ &\quad - (1-\pi) \left[s_4^* P_2^\theta + \pi(1-\pi) \right] \frac{g(s_4^*)}{X_0} \frac{\partial s_4^*}{\partial \log W_1}\end{aligned}$$

Note that the first three terms in \mathcal{E}_{11}^p are positive, while only the first two terms in \mathcal{E}_{21}^p are positive. Moreover, the second term in \mathcal{E}_{11}^p is larger than the corresponding term in \mathcal{E}_{21}^p . Overall, this implies $\mathcal{E}_{11}^p > \mathcal{E}_{21}^p$.

Similarly, we have

$$\begin{aligned}
\mathcal{E}_{21}^p &= -(1-\pi) \left[\pi f(s_0^*) (\tau W_1)^\theta + 1 \right] \frac{g(s_0^*)}{X_0} \frac{\partial s_0^*}{\partial \log P_2} + \\
&\quad - \pi^2 (1-\pi) \frac{g(s_2^*)}{X_0} \frac{\partial s_2^*}{\partial \log P_2} + \\
&\quad \pi^2 f(s_3^*) (\tau W_1)^\theta \frac{g(s_3^*)}{X_0} \frac{\partial s_3^*}{\partial \log P_2} + \\
&\quad \pi (1-\pi) \left[f(s_4^*) (\tau W_1)^\theta - (1-\pi) \right] \frac{g(s_4^*)}{X_0} \frac{\partial s_4^*}{\partial \log P_2} \\
\mathcal{E}_{22}^p &= -(1-\pi) \left[\pi (1-\pi) s_0^* P_2^\theta + 1 \right] \frac{g(s_0^*)}{X_0} \frac{\partial s_0^*}{\partial \log P_2} + \\
&\quad - \pi^2 (1-\pi) \left[s_2^* P_2^\theta + 1 \right] \frac{g(s_2^*)}{X_0} \frac{\partial s_2^*}{\partial \log P_2} + \\
&\quad - \pi^2 s_3^* P_2^\theta \frac{g(s_3^*)}{X_0} \frac{\partial s_3^*}{\partial \log P_2} + \\
&\quad - (1-\pi) \left[s_4^* P_2^\theta + \pi (1-\pi) \right] \frac{g(s_4^*)}{X_0} \frac{\partial s_4^*}{\partial \log P_2}
\end{aligned}$$

All terms of \mathcal{E}_{22}^p are positive, while the third term of \mathcal{E}_{12}^p is negative, and the fourth term of \mathcal{E}_{12}^p is smaller in absolute value than the corresponding term of \mathcal{E}_{22}^p . Moreover, the second term is larger in \mathcal{E}_{22}^p than in \mathcal{E}_{12}^p . Overall, this implies $\mathcal{E}_{22}^p > \mathcal{E}_{12}^p$.

Reallocation elasticities instead are given by

$$\begin{aligned}
\mathcal{E}_1^\tau &= -(1-\pi) \left[\pi f(s_0^*) (\tau W_1)^\theta + 1 \right] \frac{g(s_0^*)}{X_0} \frac{\partial s_0^*}{\partial \log \tau} + \\
&\quad - \pi^2 \left[f(s_1^*) (\tau W_1)^\theta + 1 \right] \frac{g(s_1^*)}{X_0} \frac{\partial s_1^*}{\partial \log \tau} + \\
&\quad \pi^2 f(s_3^*) (\tau W_1)^\theta \frac{g(s_3^*)}{X_0} \frac{\partial s_3^*}{\partial \log \tau} + \\
&\quad \pi (1-\pi) \left[f(s_4^*) (\tau W_1)^\theta - (1-\pi) \right] \frac{g(s_4^*)}{X_0} \frac{\partial s_4^*}{\partial \log \tau} \\
\mathcal{E}_2^\tau &= -(1-\pi) \left[\pi (1-\pi) s_0^* P_2^\theta + 1 \right] \frac{g(s_0^*)}{X_0} \frac{\partial s_0^*}{\partial \log \tau} + \\
&\quad - \pi^2 \frac{g(s_1^*)}{X_0} \frac{\partial s_1^*}{\partial \log \tau} + \\
&\quad - \pi^2 s_3^* P_2^\theta \frac{g(s_3^*)}{X_0} \frac{\partial s_3^*}{\partial \log \tau} + \\
&\quad - (1-\pi) \left[s_4^* P_2^\theta + \pi (1-\pi) \right] \frac{g(s_4^*)}{X_0} \frac{\partial s_4^*}{\partial \log \tau}
\end{aligned}$$

Note that the first two terms of both \mathcal{E}_1^τ and \mathcal{E}_2^τ have opposite sign, and likewise the last two terms. These terms likely offset each other if the skill density around s_0^* and s_1^* is similar, and likewise around s_3^* and s_4^* . When this is the case, the reallocation elasticity \mathcal{E}^τ is close to 0.

Outsourcing shares are given by

$$\begin{aligned}
\frac{X_0^O}{X_0} &= \tau \frac{G(s_0^*)}{X_0} \\
\frac{X_2^O}{X_2} &= \tau \frac{\mathbb{E}(s_2 | s_2 > s_4^*) (1 - G(s_4^*))}{X_2} = \tau P_2^\theta \frac{\mathbb{E}(s_2 | s_2 > s_4^*) (1 - G(s_4^*))}{X_0}
\end{aligned}$$

If the skill density is higher around s_0^* than s_4^* , the outsourcing share of input 0 is larger than that of input 2.

Proofs for Section 5

Section 5.3 states that the price of engineering services P_E and the outsourcing share in manufacturing η_{HR}^O are increasing in the productivity of the tech sector A_T , for $\theta > \max\{\sigma, 1\}$. I provide a formal proof below, studying separately each of the three stages discussed in Section 5.3.

Stage 1: $d \log P_{M,HR} = d \log P_E$, $d\eta_{HR}^{\circ} = 0$. Therefore

$$\begin{cases} \left(\epsilon \frac{\pi Q_{M,HR}}{E} + \sigma \frac{Q_{M,E}}{E} + \frac{Y_T}{A_T E} (\theta (1 - \omega_{ME}) - \sigma \omega_{ME}) \right) \alpha_{LS} d \log P_E = \frac{Y_T}{A_T E} (\theta - 1) d \log A_T \\ d \log P_{HR} = \left(1 + \frac{\sigma}{\theta} \alpha_{LS} \right) \omega_{ME} d \log P_E \end{cases}$$

From the first equation we see that P_E is increasing in A_T for $\theta > 1$. From the second equation we see that the price of HR also increases, as people reallocate away from manufacturing and towards HR.

Stage 2:

$$\begin{cases} \left(\epsilon \frac{\pi Q_{M,HR}}{E} + \sigma \frac{Q_{M,E}}{E} + \frac{Y_T}{A_T E} (\theta (1 - \omega_{ME}) - \sigma \omega_{ME}) \right) \alpha_{LS} d \log P_E + \frac{\pi Q_{M,HR}}{E} d\eta_{HR}^{\circ} = \frac{Y_T}{A_T E} (\theta - 1) d \log A_T \\ \frac{\pi Q_{M,HR}}{HR} d\eta_{HR}^{\circ} = \left[\frac{Q_{HR}}{HR} (\theta (1 - \omega_{ME}) - \sigma \omega_{ME}) \alpha_{LS} + \epsilon \left(\frac{\pi \xi_{HR}^{\circ} Q_{M,HR}}{HR} \alpha_{LS} + (1 - \alpha_{LS}) \right) \right] d \log P_E \end{cases}$$

From the second equation we see that $\frac{d\eta_{HR}^{\circ}}{d \log P_E} > 0$, and therefore plugging into the first equation we also have $\frac{d \log P_E}{d \log A_T} > 0$. The intuition is as follows. For $\eta_{HR}^{\circ} \in (0, 1)$, manufacturing firms must be indifferent between employing in-house engineers or outsourced HR specialists on the HR task. Therefore the relative price of these two tasks must be constant, and hence the ratio of manufacturing vs HR output must also be constant. Therefore, as the demand for engineers from the tech sector increases, η_{HR}° must increase in order to reduce the demand for engineers from the manufacturing sector.

Stage 3:

$$\begin{cases} - \left[\left[\frac{Y_T}{A_T E} \theta (1 - \omega_{ME}) + \sigma \left(\frac{Q_{M,E}}{E} - \frac{Y_T}{A_T E} \omega_{ME} \right) \right] - \epsilon \right] (1 - \alpha_{LS}) d \log P_{HR} + \left[\sigma \frac{Q_{M,E}}{E} + \frac{Y_T}{A_T E} (\theta (1 - \omega_{ME}) - \sigma \omega_{ME}) \right] d \log P_E = \frac{Y_T}{A_T E} (\theta - 1) d \log A_T \\ d \log P_{HR} = \frac{\frac{Q_{HR}}{HR} \omega_{ME} (\theta + \sigma)}{\frac{Q_{HR}}{HR} [\theta \alpha_{LS} + (\theta - \sigma) \omega_{ME} (1 - \alpha_{LS})] + \epsilon \left[\frac{\pi Q_{M,HR}}{HR} \alpha_{LS} + (1 - \alpha_{LS}) \right]} d \log P_E \end{cases}$$

From the second equation we see that P_{HR} is increasing in P_E . Combining we obtain

$$\begin{aligned} & \frac{\left[\frac{Q_{HR}}{HR} \theta [1 - \omega_{ME} (1 - \alpha_{LS})] + \epsilon \left[\frac{\pi Q_{M,HR}}{HR} \alpha_{LS} + (1 - \alpha_{LS}) \right] \right] \left(\sigma \frac{Q_{M,E}}{E} + \frac{Y_T}{A_T E} (\theta (1 - \omega_{ME}) - \sigma \omega_{ME}) \right) + \epsilon \frac{Q_{HR}}{HR} \omega_{ME} (1 - \alpha_{LS}) (\theta + \sigma)}{\frac{Q_{HR}}{HR} [\theta \alpha_{LS} + (\theta - \sigma) \omega_{ME} (1 - \alpha_{LS})] + \epsilon \left[\frac{\pi Q_{M,HR}}{HR} \alpha_{LS} + (1 - \alpha_{LS}) \right]} d \log P_E = \\ & = \frac{Y_T}{A_T E} (\theta - 1) d \log A_T \end{aligned}$$

so that also P_E is increasing in A_T .