Coase Meets Bellman: Dynamic Programming for Production Networks[1]

Tomoo Kikuchi[a], Kazuo Nishimura[b], John Stachurski[c] and Junnan Zhang[d]

December 7, 2020

Abstract. We show that, for a range of models related to production networks, competitive equilibria can be recovered as solutions to dynamic programs. Although these programs fail to be contractive, we show that they are fully tractable when matched with the right tools of analysis. These tools add analytical power and open new avenues for computation. As an illustration, we cover applications related to Coase's theory of the firm, including equilibria in linear production chains with transaction costs and other kinds of frictions, as well as equilibria in production networks with multiple partners. We show how the same techniques also extend to other equilibrium and decision problems, such as the distribution of management layers within firms and the spatial distribution of cities.

Keywords: Negative discounting; dynamic programming; production chains; production networks

JEL Classification: C61, D21, D90

1. Introduction

Production networks have grown in size and complexity over time, in line with advances in communications, supply chain management and transportation technology (see, e.g., Coe and Yeung (2015)). These large and complex production networks are sensitive to uncertainty, trade disputes, transaction costs and other frictions.

[1]We gratefully acknowledge valuable comments from Pol Antràs, Yongsung Chang, Davin Chor, Ryo Jinnai, Young Sik Kim, Yuta Takahashi and Makoto Saito and financial support from Korea University (K2009811 and K1922081), BrainKorea21 Plus (K1327408 and T192201), JSPS Grant-in-Aid for Scientific Research (23000001 and 23330063), Singapore Ministry of Education Academic Research Fund R-122-000-140-112 and Australian Research Council Award DP120100321. Email contacts: `tomookikuchi@korea.ac.kr`, `nishimura@rieb.kobe-u.ac.jp`, `john.stachurski@anu.edu.au`, `junnan.zhang@anu.edu.au`

[a] Department of Economics, Korea University; [b] Research Institute of Economics and Business Administration, Kobe University; [c, d] Research School of Economics, Australian National University

Firms routinely shift production and task allocation across networks, in order to mitigate risk or exploit new opportunities (see, e.g., Farlow (2020)). There is an ongoing need to predict how equilibria in production networks adapt and respond to shocks, in order to understand their impact on domestic employment, industry concentration, productivity and tax revenue.

One tool successfully used to compute and analyze equilibria in studies of firms and industries is dynamic programming (Bellman (1957)). Naturally enough, dynamic programming is most often used to analyze equilibria in *dynamic* models (see, e.g., Stokey and Lucas (1989)). However, dynamic programming can also be applied to study equilibria in static models, after the time parameter has been reinterpreted as an index over a list of firms or other decision making entities. For example, Garicano and Rossi-Hansberg (2006), Hsu et al. (2014), Tyazhelnikov (2019), and Antràs and De Gortari (2020) have applied these ideas in specific contexts. Our paper builds on this literature by providing a systematic way to apply the theory of dynamic programming to both production chains and production networks, as well as to a range of other static allocation problems involving firm management and economic geography.

In pursuing this research agenda, we are forced to confront a nontrivial technical hurdle: the dynamic programs most naturally mapped to the static competitive allocation problems we wish to consider are, in general, not contractive. Contractivity fails because frictions such as the transaction costs or failure probabilities in the production chain models translate into negative discount rates in the corresponding dynamic program.

When confronting these technical issues, we draw on dynamic programming methods most often associated with nonlinear recursive decision problems, where such technical problems are commonplace (see, e.g., Bertsekas (2013), Epstein and Zin (1989) or Bloise and Vailakis (2018)). In this sense, our work can be viewed as building connections between (a) the existing literature on dynamic programming for obtaining static competitive equilibria and (b) the modern theory of dynamic programming with recursive preferences.

The contributions provided by this paper fall into two parts. The first is providing a suitable theory of dynamic programming in a loss-minimization setting without contractivity. The second is applying this theory to a series of competitive equilibrium problems involving production chains, production networks and other related models. Through the application of this theory, we show how the dynamic programming tools can be used to obtain not only existence and uniqueness of equilibria, but also computational algorithms, results on comparative statics and insights into the underlying mechanisms.

Regarding the dynamic programming theory we present, the closest existing work in the economic literature is Bloise and Vailakis (2018). Our interiority conditions are slightly stronger than those imposed in that paper. At the same time, these conditions are satisfied in the applications we consider and that structure can be exploited to obtain strong results. For example, in addition to results on existence and uniqueness of fixed points of the Bellman operator, which parallel analogous results in Bloise and Vailakis (2018), we provide new results on monotonicity, convexity and differentiability of solutions, as well as a full set of optimality results linking Bellman's equation to existence and characterization of optimal solutions.[3]

In terms of application of these dynamic programming methods, we first connect to an analytical framework for analyzing allocation of tasks across firms in the presence of transaction costs provided by Coase (1937). Coase argued that, since market-based purchases from suppliers can be substituted for in-house operations, the size of business firms in free market economies must be determined by a choice of scale that equalizes the marginal cost of these two alternatives. While Coase's original theory lacked quantitative analysis and a role for market power, the principle that profit maximizing firms should equalize these two marginal costs remains a natural benchmark (see, e.g., Varian (2002)).

Kikuchi et al. (2018), Fally and Hillberry (2018) and Yu and Zhang (2019) develop Coasian models in which firms trade off coordination costs within the firm against transaction costs outside the firm. Firms enter a linear production chain and complete tasks sequentially. Kikuchi et al. (2018) use this model to analyze the relationship between downstreamness and firm size. Fally and Hillberry (2018) embed the model in an international trade setting and study how the structural parameters influence the gross-output-to-value-added ratio and countries' relative position within an optimized chain. Yu and Zhang (2019) extend some of these ideas to complex production networks where firms have many upstream partners.

[3] On a technical level, our optimality theory is related to other studies of dynamic programming where the Bellman operator fails to be a contraction, such as Martins-da Rocha and Vailakis (2010) and Rincón-Zapatero and Rodríguez-Palmero (2003). Our methods differ because even the relatively weak local contraction conditions imposed in that line of research fail in our settings. The fixed point results in this paper are related to those found in Kamihigashi et al. (2015), but here we also prove uniqueness of the fixed point, as well as connections to optimality and shape and differentiability properties.

We show that competitive equilibria in a version of the model extending Kikuchi et al. (2018) can be recovered as solutions to dynamic programming problems of the form described above, after a reinterpretation whereby size or scope of individual firms is mapped to effort over time. In the dynamic program, an agent acts to minimize a sequences of losses. The resulting allocation of effort is shown to be identical to the decentralized competitive equilibrium in the production problem. The envelope condition in the dynamic program maps to a condition in the production problem equalizing the marginal costs of in-house and external operations, thereby realizing one of the key conjectures of Coase (1937).

Closely related to the Coasian models above is the model by Antràs and De Gortari (2020), who use recursive methods to solve the problem of allocating production stages across countries in a global value chain. The trade-off between production costs and trade costs leads to fragmentation of production across borders in their model. They show that firms locate relatively downstream stages of production in relatively central locations where trade costs are lower. We show that analogous results can be obtained by an Euler type equation from dynamic programming.

We also consider a related application where transaction costs are replaced by another friction: failures in production or costly transportation, as found, for example, in Levine (2012) and Costinot et al. (2013). Like transaction costs that affect inter-firm trade, these features tend to inhibit fine-grained division of tasks across firms. In equilibrium, such costs are balanced against the gains from specialization. We show that the core ideas can also be analyzed via our framework.

The next application we consider is a network problem from the field of economic geography. We review the analysis of central place theory via a dynamic programming approach in Hsu et al. (2014), who show that the decentralized solution is recoverable by a recursive formulation, and that this solution can be obtained by iterating with what amounts to the Bellman operator of the programming problem. While the results in Hsu et al. (2014) are specific to the problem at hand, we show that analogous results can be studied using the general dynamic programming theory developed in this paper. In particular, we use that theory to generate a similar city hierarchy to Hsu et al. (2014), where the city size distribution follows the rank-size rule, as well as to gain additional insights into how a city hierarchy is formed.

We also consider a general production network problem in the spirit of Baldwin and Venables (2013), Kikuchi et al. (2018), Yu and Zhang (2019) and Tyazhelnikov (2019). In this production problem, production is not necessarily sequential and firms can have multiple upstream partners. The trade-off faced by firms is between transaction costs, which might also be understood as transportation costs, and costly span of control. We show that the dynamic programming theory developed below can be used to solve for competitive equilibria in these more general production networks. For example, one key feature of Tyazhelnikov's (2019) model is that firms choose to organize their production in large clusters to save on trade and unbundling costs. This feature is also present in our model.

In Section 5, we reexamine an influential model of knowledge organization due to Garicano (2000), which analyzes how workers are partitioned into classes such that each class is associated with a knowledge set to maximize output per capita. Optimal organizational forms are characterized by a pyramidal structure, in which workers are organized into layers with each layer smaller than the previous one. We show how the same dynamic programming theory developed for production networks can be used to solve this internal organization problem.

There is one more economic application of the dynamic programming theory contained in the paper, which we alluded to above. This application is presented as a running example in the section on dynamic programming theory (Section 2), and concerns an agent who seeks to minimize the present value of a sequence of losses over an infinite horizon, while assigning future losses *greater* weight than current losses (see, e.g., Section 2.1.2). In other words, the subjective discount rate is negative. Unlike the other applications in this paper, the negative discount problem is dynamic and concerns the actions of an individual agent. The reason a connection exists between such negative discount dynamic programs and the competitive equilibrium problems we consider in the rest of the paper, is that, in these equilibrium problems, frictions such as transaction costs, transportation costs and positive failure rates show up in the dynamic program as a negative discount rate.

As an optimization problem, infinite sequences of payoffs at negative discount rates cause technical difficulties. For example, in discrete time infinite horizon models, where the choice problem is represented as a dynamic program and rewards are bounded, the Bellman operator satisfies the conditions of Banach's contraction mapping theorem if and only if the discount factor is less than one (see, e.g., Stokey and Lucas (1989) or Bertsekas (2017)). This contractive property is, in turn, central to the theory of infinite horizon dynamic programming in the benchmark case. Nonetheless, we show that the dynamic programming methodology developed in the paper can handle negative discounting, by applying alternative fixed point methods from the theory of monotone convex operators (see, e.g., Zhang (2012)).[4]

The applications discussed above differ in many ways. There are different trade-offs that characterize each model, each of which leads to a particular endogenous structure. The trade-off between production costs and trade costs leads to fragmentation of production across borders in Tyazhelnikov (2019) and Antràs and De Gortari (2020). The trade-off between setup costs of cities and transport costs leads to city layers in Hsu (2012) and Hsu et al. (2014). Lastly, the trade-off between knowledge acquisition costs and communication costs leads to management layers in Garicano (2000) and Garicano and Rossi-Hansberg (2006). The dynamic programming theory developed below provides a unifying methodology and brings tools to bear on understanding the structure of the networks where firms, cities and managers save on trade, transport and communication costs, and concentrate production.

The remainder of this paper is structured as follows. In Section 2, we introduce a general dynamic programming theory that encompasses the negative discount dynamic programming theory, and discuss its solution. In Section 3, we connect this discussion to Coase's theory of the firm and elaborate on the relationship between our model and other related models. In Section 4 we extend our model to expand the scope of applications to more complex networks while in Section 5 we show that our model can also be used to understand organization of knowledge within a firm. Section 6 concludes. Most proofs are deferred to the appendix.

[4]While we do not dwell on the empirical relevance of the negative discount dynamic programming problem in the discussion above, it does have some independent value. For example, Thaler (1981), Loewenstein and Thaler (1989), Loewenstein and Prelec (1991) and Loewenstein and Sicherman (1991) all document separate instances of such phenomena. In an analysis of income path preferences, Loewenstein and Sicherman (1991) found that the majority of surveyed workers reported a preference for increasing wage profiles over decreasing ones, even when it was pointed out that the latter could be used to construct a dominating consumption sequence. Loewenstein and Prelec (1991) obtained similar results, stating that "sequences of outcomes that decline in value are greatly disliked, indicating a negative rate of time preference" (Loewenstein and Prelec, 1991, p. 351). Also, in terms of theory, Kocherlakota (1990) argues that negative discounting helps to explain the equity premium puzzle of Mehra and Prescott (1985). More recently, negative discounting can be seen in models studying monetary and fiscal policy in the presence of a zero lower bound for nominal interest rates (see, e.g., Christiano et al. (2011) or Hills et al. (2019)).

2. Noncontractive Dynamic Programming

First we provide a general dynamic programming framework suitable for analyzing equilibria in production networks. Throughout this section, we state results in an abstract setting and illustrate them with a fixed example related to minimizing a flow of losses under negative discounting.

2.1. **Set Up.**

Given a metric space E, let $\mathbb{R}^E$ denote the set of functions from E to $\mathbb{R}$ and let $c\mathbb{R}^E$ be all continuous functions in $\mathbb{R}^E$. Given $g, h \in \mathbb{R}^E$, we write $g \leqslant h$ if $g(x) \leqslant h(x)$ for all $x \in E$, and $\|f\| := \sup_{x \in E} |f(x)|$.

2.1.1. *An Abstract Dynamic Program.* Let X be a compact metric space, referred to as the state space. Let A be a metric space and let G be a nonempty, continuous, compact-valued correspondence from X into A. We understand $G(x)$ as the set of available actions $a \in A$ for an agent facing state x. Let $F_G := \{(x,a) : x \in X, \ a \in G(x)\}$ be all feasible state-action pairs. Let L be an *aggregator function*, which is a map from $F_G \times \mathbb{R}^X$ into $\mathbb{R}$, with the interpretation that $L(x,a,w)$ is lifetime loss associated with current state x, current action a and continuation value function w.

A pair (L, G) with these properties is referred to below as an *abstract dynamic program*. The *Bellman operator* associated with such a pair is the operator T defined by

$$(Tw)(x) = \inf_{a \in G(x)} L(x,a,w) \qquad (w \in \mathbb{R}^X, \ x \in X). \tag{1}$$

A fixed point of T in $\mathbb{R}^X$ is said to satisfy *the Bellman equation.*

2.1.2. *Example: Negative Discount DP.* To illustrate the definitions above, we consider an agent who takes action a_t in period t with current loss $\ell(a_t)$. We can interpret a_t as effort and $\ell(a_t)$ as disutility of effort. Her optimization problem is, for some $\hat{x} > 0$,

$$\min_{\{a_t\}} \sum_{t=0}^{\infty} \beta^t \ell(a_t) \quad \text{s.t. } a_t \geqslant 0 \text{ for all } t \geqslant 0 \text{ and } \sum_{t=0}^{\infty} a_t = \hat{x}. \tag{2}$$

Here and throughout the discussion of this optimization problem, we suppose that

$$\beta > 1, \ \ell(0) = 0, \ \ell' > 0 \ \text{ and } \ell'' > 0. \tag{3}$$

The convexity in ℓ encourages the agent to defer some effort. Negative discounting ($\beta > 1$) has the opposite effect.[5] We set $x_{t+1} = x_t - a_t$ with $x_0 = \hat{x}$, so x_t represents "remaining tasks" at the start of time t. The Bellman equation is

$$w(x) = \inf_{0 \leqslant a \leqslant x} \{\ell(a) + \beta w(x-a)\}. \tag{4}$$

The Bellman operator is

$$Tw(x) = \inf_{0 \leqslant a \leqslant x} \{\ell(a) + \beta w(x-a)\}. \tag{5}$$

If we set $X := A := [0, \hat{x}]$,

$$L(x, a, w) := \ell(a) + \beta w(x-a) \quad \text{and} \quad G(x) := [0, x], \tag{ND}$$

then (L, G) in (ND) fits the definition of an abstract dynamic program, as given in Section 2.1.1, and (5) is a special case of (1).

2.2. **Fixed Point Results.** The Bellman operator in (5) is not a supremum norm contraction, due to the fact that $\beta > 1$.[6] The production chain and network models we consider also have this feature. Hence we introduce a set of conditions for an abstract dynamic program that generate stability properties without requiring contractivity.

2.2.1. *Stability Without Contractivity.* Fix an abstract dynamic program (L, G) and consider the following assumptions:

A_1. $(x, a) \mapsto L(x, a, w)$ is continuous on F_G when $w \in c\mathbb{R}^X$.
A_2. If $u, v \in c\mathbb{R}^X$ with $u \leqslant v$, then $L(x, a, u) \leqslant L(x, a, v)$ for all $(x, a) \in F_G$.
A_3. Given $\lambda \in (0, 1)$, $u, v \in c\mathbb{R}^X$ and $(x, a) \in F_G$, we have

$$\lambda L(x, a, u) + (1-\lambda) L(x, a, v) \leqslant L(x, a, \lambda u + (1-\lambda) v).$$

[5]The assumption $\ell(0) = 0$ cannot be weakened, since $\ell(0) > 0$ implies that the objective function is infinite. Conversely, with the assumption $\ell(0) = 0$, minimal loss is always finite. Indeed, by choosing the feasible action path $a_0 = \hat{x}$ and $a_t = 0$ for all $t \geqslant 1$, we get $\sum_{t=0}^{\infty} \beta^t \ell(a_t) \leqslant \ell(\hat{x})$. Also, given our other assumptions, there is no need to consider the case $\beta \leqslant 1$ because no solution exists. Because we are minimizing loss, when $\beta < 1$ any proposed solution $\{a_t\}$ can be strictly improved by shifting it one step into the future (set $a'_0 = 0$ and $a'_{t+1} = a_t$ for all $t \geqslant 0$). Furthermore, if $\beta = 1$, and a solution $\{a_t\}$ exists, then the increments $\{a_t\}$ must converge to zero, and hence there exists a pair a_T and a_{T+1} with $a_T > a_{T+1}$. Since ℓ is strictly convex, the objective $\sum_t \ell(a_t)$ can be reduced by redistributing a small amount ε from a_T to a_{T+1}. This contradicts optimality.

[6]For example, let $w \equiv 1$ and $g \equiv 0$. Then $Tw \equiv \beta > 1$ while $Tg \equiv 0$. One consequence is that, if we take an arbitrary continuous bounded function and iterate with T, the sequence typically diverges. For example, if $w \equiv 1$, then, $T^n w \equiv \beta^n$, which diverges to $+\infty$.

A_4. There is a ψ in $c\mathbb{R}^X$ such that $T\psi \leqslant \psi$.

A_5. There is a φ in $c\mathbb{R}^X$ and an $\varepsilon > 0$ such that $\varphi \leqslant \psi$ and $T\varphi \geqslant \varphi + \varepsilon(\psi - \varphi)$.

Assumptions A_1–A_3 impose some continuity, monotonicity and convexity. Assumptions A_4–A_5 provide upper and lower bounds for the set of candidate value functions.

Although contractivity is not imposed, we can show that the abstract Bellman operator (1) is well behaved under A_1–A_5 after restricting its domain to a suitable class of candidate solutions. To this end, let

$$\mathcal{I} := \{f \in c\mathbb{R}^X : \varphi \leqslant f \leqslant \psi\}.$$

Theorem 2.1. *Let (L, G) be an abstract dynamic program and let T be the Bellman operator defined in (1). If (L, G) satisfies A_1–A_5, then*

1. *T has a unique fixed point w^* in $\mathcal{I}$.*
2. *For each $w \in \mathcal{I}$, there exists an $\alpha < 1$ and $M < \infty$ such that*
$$\|T^n w - w^*\| \leqslant \alpha^n M \quad \textit{for all } n \in \mathbb{N}. \tag{6}$$
3. *$\pi^*(x) := \arg\min_{a \in G(x)} L(x, a, w^*)$ is upper hemicontinuous on X.*

Theorem 2.1 does not discuss Bellman's principle of optimality. That task is left until Section 2.4. Regarding π^*, which has the interpretation of a policy correspondence, an immediate corollary is that π^* is continuous whenever π^* is single-valued on X.

2.2.2. *Example: Negative Discount DP.* Let (L, G) be defined by (ND) and suppose the assumptions on β and ℓ in (3) are true. Let $\varphi(x) := \ell'(0)x$ and $\psi(x) := \ell(x)$ be the boundary functions in A_4–A_5. Assumptions A_1–A_3 clearly hold. The condition $T\psi \leqslant \psi$ in A_4 also holds because, with $\psi = \ell$,

$$T\psi(x) = T\ell(x) = \inf_{0 \leqslant a \leqslant x} \{\ell(a) + \beta\ell(x - a)\} \leqslant \ell(x) + \beta\ell(0) = \ell(x). \tag{7}$$

Under the auxiliary assumption $\ell'(0) > 0$, we show that A_5 also holds. The details are in the appendix (see Proposition 7.5). Hence, under these assumptions, the conclusions of Theorem 2.1 are valid for the Bellman operator T defined in (5).

2.3. **Shape and Smoothness Properties.** We now give conditions under which the solution to the Bellman equation associated with an abstract dynamic program possesses additional properties, including monotonicity, convexity and differentiability. In what follows, we assume that X is convex in $\mathbb{R}$ and F_G is convex in $X \times A$. We let

1. $ic\mathbb{R}^X$ be all increasing functions in $c\mathbb{R}^X$ and
2. $cc\mathbb{R}^X$ be all convex functions in $c\mathbb{R}^X$.

We assume that $\mathcal{I}$ defined above contains at least one element of each set.

2.3.1. *Results.* To obtain convexity and differentiability, we impose

Assumption 2.1. In addition to A_1–A_5, the abstract dynamic program (L, G) satisfies the following conditions:

1. If $w \in cc\mathbb{R}^X$, then $(x, a) \to L(x, a, w)$ is strictly convex on F_G.
2. If $a \in \operatorname{int} G(x)$ and $w \in cc\mathbb{R}^X$, then $x \to L(x, a, w)$ is differentiable on $\operatorname{int} X$.

We can now state the following result.

Theorem 2.2. *If Tw is strictly increasing for all $w \in ic\mathbb{R}^X$, then w^* is strictly increasing. If Assumption 2.1 holds, then w^* is strictly convex, π^* is single-valued, w^* is differentiable on* $\operatorname{int} X$ *and*

$$(w^*)'(x) = L_x(x, \pi^*(x), w^*) \tag{8}$$

whenever $\pi^*(x) \in \operatorname{int} G(x)$.

2.3.2. *Example: Negative Discount DP.* Using Theorem 2.2, we can derive properties of the negative discount dynamic program defined in (ND). In particular, we can show that the fixed point w^* of the negative discount Bellman operator T in (5) is strictly increasing, strictly convex, and continuously differentiable on $(0, \hat{x})$, and that π^* is single-valued and satisfies the envelope condition

$$(w^*)'(x) = \ell'(\pi^*(x)) \qquad (0 < x < \hat{x}). \tag{9}$$

To obtain (9) from (8), we use the change of variable $y = x - a$ to write

$$w^*(x) = \min_{0 \leqslant a \leqslant x} \{\ell(a) + \beta w^*(x - a)\} = \min_{0 \leqslant y \leqslant x} \{\ell(x - y) + \beta w^*(y)\}.$$

Differentiating the final term with respect to x and evaluating at the optimal choice gives (9).

We provide a more detailed proof of (9) and proofs of other claims from this section in Proposition 7.6 in the appendix. We rely on the convexity and differentiability of ℓ to check Assumption 2.1.

2.4. **The Principle of Optimality.** If we consider the implications of the preceding dynamic programming theory, we have obtained existence of a unique solution to the Bellman equation and certain other properties, but we still lack a definition of optimal policies, and a set of results that connect optimality and solutions to the Bellman equation. This section fills these gaps.

Let Π be all $\pi \colon X \to A$ such that $\pi(x) \in G(x)$ for all $x \in X$. For each $\pi \in \Pi$ and $w \in \mathbb{R}^X$, define the operator T_π by

$$(T_\pi w)(x) = L(x, \pi(x), w). \tag{10}$$

This can be understood as the lifetime loss of an agent following π with continuation value w. Let $\mathcal{M}$ be the set of *(nonstationary) policies*, defined as all $\mu = \{\pi_0, \pi_1, \dots\}$ such that $\pi_t \in \Pi$ for all t. For stationary policy $\{\pi, \pi, \dots\}$, we simply refer it as π. Let the *μ-value function* be defined as

$$w_\mu(x) := \limsup_{n \to \infty} (T_{\pi_0} T_{\pi_1} \dots T_{\pi_n} \varphi)(x), \tag{11}$$

where φ is the lower bound function in $\mathcal{I}$. Note that w_μ is always well defined. The agent's problem is to minimize w_μ by choosing a policy in $\mathcal{M}$. The *value function* $\bar{w}$ is defined by

$$\bar{w}(x) := \inf_{\mu \in \mathcal{M}} w_\mu(x) \tag{12}$$

and the *optimal policy* $\bar{\mu}$ is such that $\bar{w} = w_{\bar{\mu}}$. We impose the following assumption.

Assumption 2.2. In addition to A_1–A_5, the abstract dynamic program (L, G) satisfies the following conditions:

1. If $(x, a) \in F_G$, $v_n \geqslant \varphi$ and $v_n \uparrow v$, then $L(x, a, v_n) \to L(x, a, v)$.
2. There exists a $\beta > 0$ such that, for all $(x, a) \in F_G$, $r > 0$ and $w \geqslant \varphi$,

$$L(x, a, w + r) \leqslant L(x, a, w) + \beta r. \tag{13}$$

Part 1 of Assumption 2.2 is a weak continuity requirement on the aggregator with respect to the continuation value, similar to Assumption 4 in Bloise and Vailakis (2018). Part 2 of Assumption 2.2 is analogous to the Blackwell's condition, with the significant exception that β in (13) is not restricted to be less than one.

Theorem 2.3. *If Assumption 2.2 holds, then $w^* = \bar{w}$ and an optimal stationary policy exists. Moreover, a stationary policy π is optimal if and only if $T_\pi \bar{w} = T\bar{w}$.*

Theorem 2.3 shows that the fixed point of the Bellman operator is the value function and the Bellman's principle of optimality holds. It immediately follows that any selector of π^* in Theorem 2.1 is an optimal stationary policy.

2.4.1. *Example: Negative Discount DP.* Consider again the negative discount dynamic program (L, G) defined in (ND), under the assumptions in (3). Both conditions in Assumption 2.2 can be verified for (L, G). Part 1 of Assumption 2.2 is trivial in this setting, since $v_n \uparrow v$ pointwise clearly implies $\ell(a) + \beta v_n(x-a) \to \ell(a) + \beta v(x-a)$ at each $(x, a) \in F_G$. Part 2 also holds, since for any $r > 0$ and $w \geqslant \varphi$, we have

$$L(x, a, w + r) = \ell(a) + \beta w(x - a) + \beta r = L(x, a, w) + \beta r.$$

Hence Theorem 2.3 applies. In fact, in this setting we can be more explicit, by setting

$$W(x) := \min \left\{ \sum_{t=0}^{\infty} \beta^t \ell(a_t) \;\; : \;\; \{a_t\} \in \mathbb{R}_+^\infty \text{ and } \sum_{t=0}^{\infty} a_t = x \right\} \tag{14}$$

at each $x \geqslant 0$. By construction, $W(\hat{x})$ is the minimum cost in (2). To connect W and the fixed point w^*, we first show that (14) is equivalent to (12) and then apply Theorem 2.3. The details are in Proposition 7.9 in the appendix, which shows that W is the solution to the Bellman equation (4), the principle of optimality holds, and there exists a unique solution to (2) given by $a_t^* = \pi^*(x_t)$, where the state process is governed by $x_{t+1} = x_t - a_t^*$ and $x_0 = \hat{x}$.

The envelope condition (9) now evaluates to

$$W'(x_t) = \ell'(a_t^*) \tag{EN}$$

for all $t \in \mathbb{Z}$, which links marginal value to marginal disutility at optimal action. Furthermore, (EN) implies that the sequence $\{a_t^*\}$ satisfies[7]

$$\ell'(a_{t+1}^*) = \max \left\{ \frac{1}{\beta} \ell'(a_t^*), \; \ell'(0) \right\} \tag{EU}$$

for all $t \in \mathbb{Z}$, which is akin to an Euler equation with a possibly binding constraint. In the applications below we use (EN) and (EU) to aid interpretation and provide economic intuition.

It follows immediately from (EU) that $\{a_t^*\}$ is a decreasing sequence. This concurs with our intuition: future losses are given greater weight than current losses, so $\{a_t^*\}$ declines over time.

[7]To see this, note that a_t^* solves $\inf_{0 \leqslant a \leqslant x_t} \{\ell(a) + \beta w^*(x_t - a)\}$. Since both ℓ and w^* are convex, elementary arguments show that either $\ell'(a_t^*) = \beta (w^*)'(x_t - a_t^*)$ or $a_t^* = x_t$. It follows from (EN) that either $\ell'(a_t^*) = \beta \ell'(a_{t+1}^*)$ or $a_{t+1}^* = 0$, which is equivalent to (EU).

2.5. **Additional Results.** Some additional results hold for the specific case of the negative discount dynamic program introduced in Section 2.1.2. One result is a strong form of convergence for the Bellman operator T from (5). In particular, iteration always converges in finite time. The details are in Proposition 7.10 in the appendix. In addition, we can treat the case $\ell'(0) = 0$, which has hitherto been excluded:

Proposition 2.4. *When $\ell'(0) = 0$, a feasible sequence $\{a_t^*\}$ solves (2) if and only if (EU) holds. This sequence is unique, decreasing, and satisfies $a_t^* > 0$ for all t.*

Proposition 2.4 shows that the Euler equation (EU) established above becomes a necessary and sufficient condition for optimality in this case. In fact, (EU) can be reduced to $\beta\ell'(a_{t+1}^*) = \ell'(a_t^*)$ when $\ell'(0) = 0$, which helps us derive analytical solutions for some of the applications below.

As the above results suggest, the set of tasks will be completed in finite time if and only if $\ell'(0) > 0$. When the agent never finishes in finite time, the corner solution $a_t^* = 0$ never binds, and the optimality result in Proposition 2.4 can be established through elementary arguments. The proof is in the appendix.

3. Application: Production Chains

Now we turn to applications of Theorem 2.1 motivated by production problems. We begin with linear production chains.

3.1. **A Coasian Production Chain.** The size distribution of business firms has significant impact on macroeconomic volatility (see, e.g., Carvalho and Grassi (2019)) and is a recurring topic in policy discussions. Early contributions to our understanding of the determinants of firm size were made by Coase (1937). He argued that, on one hand, in-house production becomes expensive as the scope of operations increases due to bureaucratic inefficiencies.[8] On the other hand, free market purchases from suppliers attract transaction costs. Firms size reflects an attempt to equalize the marginal costs of in-house and market-based operations.

Quantitative expressions of the these ideas have been put forward by Kikuchi et al. (2018), Fally and Hillberry (2018) and Yu and Zhang (2019) in the context of competitive equilibrium models of production chains. We consider a version of these

[8]These ideas appear in many studies under different names. For example, it finds expression in the "span-of-control" costs highlighted by Lucas (1978), Buera and Shin (2011) or Clementi and Palazzo (2016), and the "coordination costs" of Becker and Murphy (1992).

models and show how the competitive equilibrium can be calculated using the dynamic programming theory from Section 2. This permits weaker assumptions and generates additional insights.

3.1.1. *Set Up.* Consider a large and competitive market with many price-taking firms, each of which is either inactive or involved in the production of a single good. To produce a unit of this good requires implementing a sequence of tasks, indexed by $s \in [0, 1]$. A production chain is a collection of firms that implements all of these tasks and produces the final good. Firms face no fixed costs or barriers to entry.

Let $c(v)$ be the cost for any one firm of implementing an interval of tasks with length v. We assume that c is increasing, strictly convex, continuously differentiable, and satisfies $c(0) = 0$. (Unlike Kikuchi et al. (2018), we allow $c'(0) = 0$.) Firms face transaction costs, as a wedge between price paid by the buyer and payment received by the seller.[9] For convenience, we assume that the transaction cost falls entirely on the buyers, so that, for a transaction with face value f, the seller receives f and the buyer pays $(1 + \tau)f$, where $\tau > 0$.[10]

Firms are indexed by integers $i \geqslant 0$. A feasible allocation of tasks across firms is a nonnegative sequence $\{v_i\}$ with $\sum_{i \geqslant 0} v_i = 1$. We identify firm 0 with the most downstream firm, firm 1 with the second most downstream firm, and so on. Let b_i be the downstream boundary of firm i, so that $b_0 = 1$ and $b_{i+1} = b_i - v_i$ for all $i \in \mathbb{Z}$. Then, profits of the ith firm are

$$\pi_i = p(b_i) - c(v_i) - (1 + \tau)p(b_{i+1}). \tag{15}$$

Here $p\colon [0, 1] \to \mathbb{R}_+$ is a price function, with $p(t)$ interpreted as the price of the good at processing stage t.

Definition 3.1. Given a price function p and a feasible allocation $\{v_i\}$, let $\{\pi_i\}$ be corresponding profits, as defined in (15). The pair $(p, \{v_i\})$ is called an *equilibrium* for the production chain if

1. $p(0) = 0$,
2. $p(s) - c(s - t) - (1 + \tau)p(t) \leqslant 0$ for any pair s, t with $0 \leqslant t \leqslant s \leqslant 1$, and

[9] This follows Kikuchi et al. (2018) and also studies such as Boehm and Oberfield (2018), where frictions in contract enforcement are treated as a variable wedge between effective cost to the buyer and payment to the supplier.

[10] For example, τf might be the cost of writing a contract for a transaction with face value f. This cost rises in f because more expensive transactions merit more careful contracts. (There are other possible interpretations for τ, some of which are touched on below.)

3. $\pi_i = 0$ for all i.

Condition 1 rules out profits for suppliers of initial inputs, which are assumed for convenience to have zero cost of production. Condition 2 ensures that no firm in the production chain has an incentive to deviate, and that inactive firms cannot enter and extract positive profits. Condition 3 requires that active firms make zero profits, due to free entry and an infinite fringe of potential competitors.

3.1.2. *Solution by Dynamic Programming.* Note that an equilibrium of the production chain satisfies $p(b_i) = c(v_i) + (1+\tau)p(b_i - v_i)$, which has the same form as the Bellman equation (4). Moreover, iterating on this relation yields the price of the final good

$$p(1) = \sum_{i \geqslant 0} (1+\tau)^i c(v_i), \tag{16}$$

which is analogous to (2). These facts motivate us to consider a version of the negative discount dynamic program introduced in Section 2.1.2 where a (fictitious) agent seeks to minimize $\sum_{i \geqslant 0}(1+\tau)^i c(a_i)$ subject to $\sum_{i \geqslant 0} a_i = 1$. In other words, we specialize the problem to one where $\hat{x} = 1$, $\ell = c$ and $\beta = 1+\tau$. By construction, any feasible action path is also a feasible allocation of tasks in the production chain.

Since the assumptions in Section 2.1.2 are satisfied, we know that there exists a unique solution $\{a_i^*\}$. Let W be the corresponding value function given by (14). The next proposition shows that the solution to this dynamic program is precisely the competitive equilibrium of the Coasian production chain described above. In view of (16), it follows that the equilibrium allocation also gives the minimum price for the final good.

Proposition 3.1. *Let $p = W$ and $v_i = a_i^*$ for all $i \in \mathbb{Z}$. Then the pair $(p, \{v_i\})$ is an equilibrium for the production chain.*

One insight from this result is as follows. We know from Section 2.3.2 that the price function is continuously differentiable on $(0,1)$ and, for firm with downstream boundary b_i,

$$p'(b_i) = c'(v_i), \tag{17}$$

which follows from the envelope condition (EN). Since v_i is the optimal range of tasks implemented in-house by firm i in equilibrium, this is an expression of Coase's key idea: the size of the firm is determined as the scale that equalizes the marginal costs of in-house and market-based operations. This trade-off is further clarified in the Euler equation (EU), which says that each firm achieves optimum when the cost of implementing an additional task equals the cost of purchasing it from a supplier. The Euler equation (EU) also implies that $\{v_i\}$ is decreasing. In other words, firm size increases with downstreamness. This generalizes a finding of Kikuchi et al. (2018).

3.1.3. *An Example.* Suppose that the range of tasks v implemented by a given firm satisfies $v = f(k, n)$, where k is capital and n is labor. Given rental rate r and wage rate w, the cost function is $c(v) := \min_{k,n}\{rk + wn\}$ subject to $f(k, n) \geqslant v$. Let us suppose further that, as in Lucas (1978), the production function has the form $\varphi(g(k, n))$, where g has constant returns to scale and φ is increasing and strictly concave, with the latter property due to "span-of-control" costs. To generate a closed-form solution, we take $g(k, n) = Ak^\alpha n^{(1-\alpha)}$ and $\varphi(x) = x^\eta$, with $0 < \alpha, \eta < 1$. The resulting cost function has the form $c(v) = \kappa v^{1/\eta}$, where κ is a positive constant.

By Proposition 3.1, the optimal action path for the fictitious agent corresponds to the equilibrium allocation of tasks across firms, and the value function is the equilibrium price function. Since $c'(0) = 0$, Proposition 2.4 applies and the Euler equation (EU) yields $a^*_{i+1} = \theta a^*_i$ for all $i \in \mathbb{Z}$, where $\theta := (1 + \tau)^{\eta/(\eta-1)} < 1$. From $\sum_{i=0}^\infty a^*_i = 1$ we obtain $v_i = a^*_i = \theta^i(1 - \theta)$. Substituting this path into (14) gives the price function

$$p(x) = W(x) = \kappa \, (1 - \theta)^{(1-\eta)/\eta} \, x^{1/\eta}. \tag{18}$$

As anticipated by the theory, p is strictly increasing and strictly convex.

Although this example lies outside the framework of Kikuchi et al. (2018), since $c'(0) = 0$, we have replicated some of their key results. For example, we have found that the size of firms increases from upstream to downstream (recall that upstream firms have larger i), and that the price function is strictly convex due to the costly span of control. Intuitively, firm-level span-of-control costs cannot be eliminated in aggregate due to transaction costs, which force firms to maintain a certain size. This leads to strict convexity of prices. If firms have constant returns to management ($\eta = 1$), then the price function in (18) becomes linear.

The above result on the size of firms is related to the main result in Antràs and De Gortari (2020), who show that it is optimal to locate relatively downstream stages of production in relatively central locations where trade costs are lower as firms save on trade costs and concentrate their production. Their result holds because trade costs have more pronounced effects in more downstream stages of production in their model. Similarly, in our model, transaction costs have more pronounced effects in more downstream states of production, which is a direct consequence of the Euler equation (EU).

3.2. **Specialization and Failure Probabilities.** Production processes typically consist of a series of complementary tasks, and mistakes in any of the tasks can dramatically reduce the product's value. Implications of such specialization and failure probabilities were studied in, among others, the O-ring theory of economic development by Kremer (1993) and the production chain models of Levine (2012) and Costinot et al. (2013). To mitigate the exponential impact on the product's value, Kremer's O-ring theory has an assortative matching of workers who have different probabilities of making mistakes and the length of production chains (number of tasks) adjusts accordingly. For example, high failure workers are grouped together and build a shorter production chain. The models of Levine (2012) and Costinot et al. (2013) have similar features, where equilibrium allocations serve to mitigate the potentially exponential cost of failures in long production chains.[11] In this section, we show that the ideas in Kremer (1993), Levine (2012) and Costinot et al. (2013) are also amenable to analysis using the dynamic programming theory from Section 2.

Consider, as before, a competitive market where producers implement a sequence of tasks indexed by $s \in [0, 1]$. We remove the assumption of positive transaction costs. Instead, the friction between firms is due to positive probability of defects. (Defects can alternatively be understood as iceberg costs, where some percentage of goods are lost in transporting them from one producer to the next.) Due to these defects, a producer who buys at stage t and sells at $s > t$ must buy $1 + \tau$ units of the partially completed good at t to sell one unit of the processed good at s. Larger τ then corresponds to a production process that is more prone to failure. Profits for such a firm when confronting price function p are

$$\pi = p(s) - c(s - t) - (1 + \tau)p(t).$$

This parallels the profit function (15) from the Coasian production chain model and the rest of the analysis is essentially identical. In particular, if we adopt the Cobb–Douglass production technology from Section 3.1.3, then the price of the final good

[11]In Levine (2012), long chains involve a high degree of specialization and produce a large quantity of output but are also more prone to failure. However, chains in his model are long only if the failure rate is low thus mitigating the exponential impact that production failure of a single link has on output. Similarly, Costinot et al. (2013), in a global supply chain model where production of the final goods is sequential and subject to mistakes, show that countries with lower probabilities of making mistakes specialize in later stages of production.

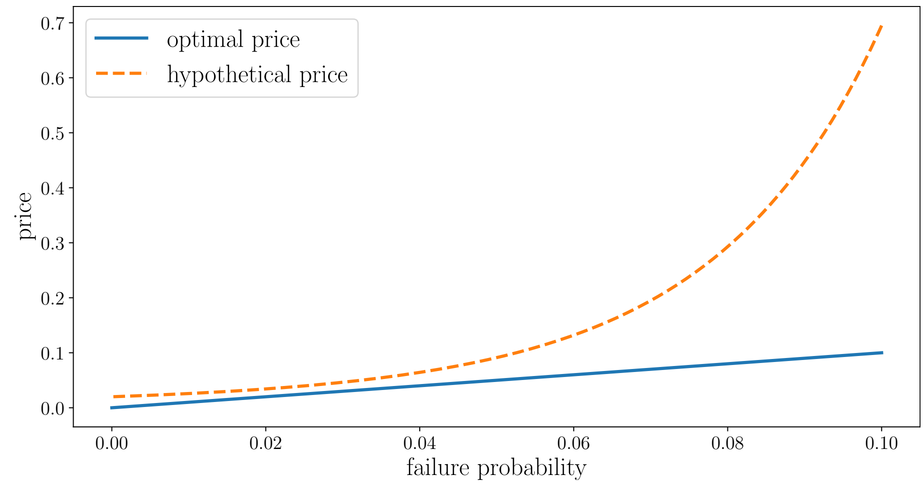


FIGURE 1. Final good price and failure probabilities.

is

$$p^*(1) = \kappa \left(1 - (1+\tau)^{\eta/(\eta-1)}\right)^{(1-\eta)/\eta}. \tag{19}$$

In the expression for $p^*(1)$ in (19), one interesting feature is that a rise in the failure probability leads to only a moderate increase in the final good price. This is because producers increase their range of internal production to mitigate any rise in cost associated with a higher production failure of upstream producers. As a result, there are fewer producers in production chains and the compounding effect of higher production failures is limited.

To clarify this point, let us compare this outcome with a hypothetical model where producers do not adjust their production according to failure probabilities. Suppose in particular that production chains are simply divided into equal tasks by N producers. In this case, the final good price is

$$\hat{p}^*(1) = \kappa \sum_{i=0}^{N} (1+\tau)^i \left(\frac{1}{N}\right)^{1/\eta} = \kappa \frac{(1+\tau)^N - 1}{(1+\tau) - 1} \left(\frac{1}{N}\right)^{1/\eta} = O((1+\tau)^N). \tag{20}$$

Now a small increase in the failure probability increases the final good price exponentially. This is intuitive, as an increase in cost compounds over all producers involved in the production chain. See Figure 1 for a comparison[12] of prices with and without producers adjusting for failure probabilities.

Thus, returning to the original model, we see that equilibrium prices induce producers to adjust to changes in failure probabilities, which optimally mitigates the

[12] In this example, we set $\kappa = 1$, $\eta = 0.5$, and $N = 50$.

potentially exponential impact of failures on the cost of the final good. This outcome is similar to the results obtained by Levine (2012) and Costinot et al. (2013) as discussed above.

4. Application: Networks

In this section we treat more general network models. Unlike the linear production chains discussed above, agents can interact with multiple partners. As before, our objective is to apply the dynamic programming theory developed in Section 2.

4.1. **Spatial Networks.** The distribution of city sizes shows remarkable regularity, as described by the rank-size rule.[13] One early attempt to match the empirical city size distribution is found in the central place theory of Christaller (1933). Hsu (2012) formalizes Christaller's theory in a model where a city hierarchy arises as a market equilibrium, while Hsu et al. (2014) shows that the market equilibrium allocation is identical to the social planner's solution. In this section, we show how a model similar to that of Hsu (2012) can be studied using the dynamic programming theory from Sections 2. We then use the Euler equation and envelop condition to gain insights into how a city hierarchy is formed.

Consider a government that opens competition for many developers to build cities to host a continuum of dwellers of measure one. One developer can build a large city that hosts everyone or build a smaller city and assign other developers to build "satellite cities" that host the rest of the population. Further satellites can be built for existing cities until all dwellers are accommodated. This chain of building layers of cities starts with a single developer, who is assigned the whole population. Building satellite cities incurs extra costs that are charged as an ad valorem tax on the payments to the developers. We can think of the extra costs as costs of providing public goods that connect different cities such as roads, electricity, water, telecommunication, etc.

Let developers be paid according to a price $p\colon [0,1] \to \mathbb{R}$, which is a function of the population assigned. Let the cost function of building and expanding a city be $c\colon [0,1] \to \mathbb{R}$ and the tax rate be τ. Then, a developer assigned to host s dwellers maximizes profits by solving

$$\max_{0 \leqslant t \leqslant s} \left\{ p(s) - c(s-t) - (1+\tau) k p(t/k) \right\},$$

[13]See Gabaix and Ioannides (2004) and Gabaix (2009) for surveys.

where $p(s)$ is the payment to the developer, $c(s-t)$ is the cost of building a city of population $s-t$, k is the number of satellite cities, and $(1+\tau)kp(t/k)$ is the cost of assigning population t/k to k satellites. In equilibrium, a city network is formed where every dweller is accommodated and every developer makes zero profits. The equilibrium price function satisfies

$$p(s) = \min_{0 \leqslant t \leqslant s} \{c(s-t) + (1+\tau)kp(t/k)\}, \tag{21}$$

which is a Bellman equation similar to (4) in Section 2.1.2. We let $c(s) = s^\gamma$ with $\gamma > 1$. To emulate the bifurcation process in Hsu (2012) and Hsu et al. (2014), we let $k = 2$.

We can formulate a dynamic programming problem similar to (2) and show that the value function satisfies (21). Consider now the same problem from the perspective of a social planner who minimizes the total cost of hosting the whole population with value function

$$W(x) := \min_{\{v_i\}} \left\{ \sum_{i=0}^{\infty} (1+\tau)^i k^i c(v_i) : \{v_i\} \in \mathbb{R}_+^\infty \text{ and } \sum_{i=0}^{\infty} k^i v_i = x \right\},$$

where v_i is the size of cities on layer i. Since the assumptions in Section 2.1.2 are satisfied, a similar argument to the proof of Proposition 2.4 gives the Euler equation

$$c'(v_i) = (1+\tau)c'(v_{i+1}). \tag{22}$$

Using this equation, it can be shown with some algebra that $v_i = \theta^i(1-2\theta)$ if $\theta := (1+\tau)^{1/(1-\gamma)} < 1/2$ and the value function is $W(s) = (1-2\theta)^{\gamma-1}s^\gamma$. It is straightforward to verify that $p = W$ satisfies (21). Hence, the minimum value that can be achieved is also the equilibrium price function under which no developer makes positive profits.

The Euler equation (22) describes the emergence of optimal city hierarchy where each developer expands a city to accommodate more dwellers until the marginal cost of expanding equals the marginal cost of building and expanding satellite cities. An envelope condition similar to (EN) also holds: if a developer is assigned s dwellers and delegate t dwellers to satellite cities, the equilibrium is reached when $p'(s) = c'(s-t)$. This shows that the marginal value that a city provides must be equal to the marginal cost of accommodating one more city dweller.

Figure 2 illustrates the optimal city hierarchy by placing cities according to Hsu (2012) and Hsu et al. (2014), where C_i represents a city on layer i.[14] It replicates

[14] We set $\gamma = 1.2$ and $\tau = 0.2$.

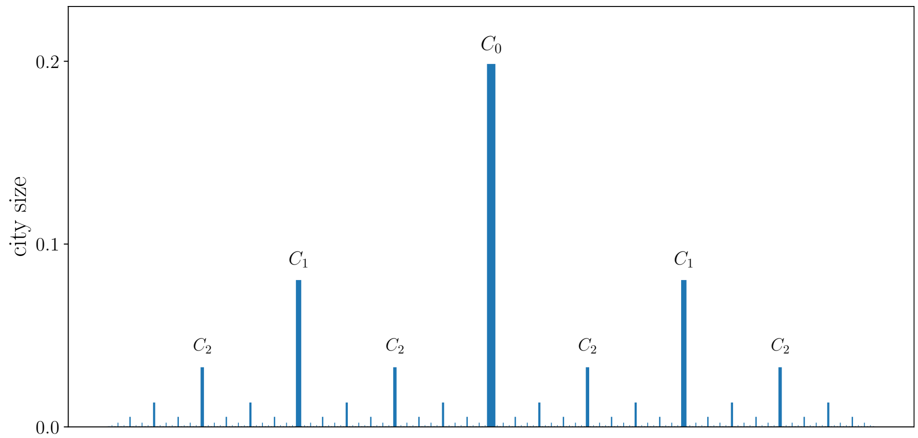


FIGURE 2. Illustration of optimal city hierarchy.

the relative sizes of cities on different layers as in Hsu (2012) and Hsu et al. (2014). Moreover, since the number of cities doubles from one layer to the next, the rank of a city on layer i is 2^i. Hence, the city size distribution generated by our model follows a power law similar to Hsu (2012). In fact, the rank and size of a city satisfy

$$\ln(Rank) = -\frac{\ln(1/2)}{\ln(\theta)} \ln(Size) + C,$$

where C is a constant determined by θ. When θ approaches 1/2, the slope approaches one, which corresponds to the well-documented rank-size rule.

4.2. **Snakes and Spiders.** Modern production networks are characterized by processes that are both sequential and non-sequential where firms assemble parts in no particular order. There are costs associated with extending each process and changes in those costs affect how production networks are formed. Baldwin and Venables (2013) refer to the sequential process as "snakes" and the non-sequential process as "spider", and analyze how the location of different parts of a production chain is determined by unbundling costs of production across borders. Here we show that the dynamic programming theory developed in Section 2 can be used to solve a general production network, which can replicate key results in the literature.

As in Kikuchi et al. (2018) and Yu and Zhang (2019), we consider a generalization of the production chain model in Section 3.1, where each firm can also choose the number of suppliers. The production chain then becomes a combination of "snakes" and "spiders". To account for costs of extending "spiders" we assume that firms bear an additive assembly cost g that is strictly increasing in the number of suppliers, with $g(1) = 0$. Then for a firm at stage s that subcontracts tasks of range t to k

suppliers, the profits are

$$p(s) - c(s-t) - g(k) - (1+\tau)kp(t/k),$$

where p is the price function. Having multiple suppliers leads to another trade-off: firms potentially benefit from subcontracting at a lower price but also have to pay additional assembly costs.

We index the layers in the production network by $i \in \mathbb{Z}$ with layer 0 consisting only of the most downstream firm. Let b_i be the downstream boundary of firms on layer i, each producing v_i and having k_i suppliers. Then the boundary of firms on the next layer is given by $b_{i+1} = (b_i - v_i)/k_i$. Similar to Definition 3.1, we call the triplet $(p, \{v_i\}, \{k_i\})$ an equilibrium for the production network if (i) $p(0) = 0$, (ii) $p(s) - c(s-t) - g(k) - (1+\tau)kp(t/k) \leqslant 0$ for all $0 \leqslant t \leqslant s \leqslant 1$ and $k \in \mathbb{N}$, and (iii) $\pi_i = 0$ for all $i \in \mathbb{Z}$ where

$$\pi_i := p(b_i) - c(v_i) - g(k_i) - (1+\tau)kp\left(\frac{b_i - v_i}{k_i}\right). \tag{23}$$

As in Section 3.1.2, we seek to find an equilibrium using dynamic programming methods. Let p^* be the solution to the following Bellman equation

$$p(s) = \min_{\substack{0 \leqslant t \leqslant s \\ k \in \mathbb{N}}} \left\{ c(s-t) + g(k) + (1+\tau)kp(t/k) \right\}. \tag{24}$$

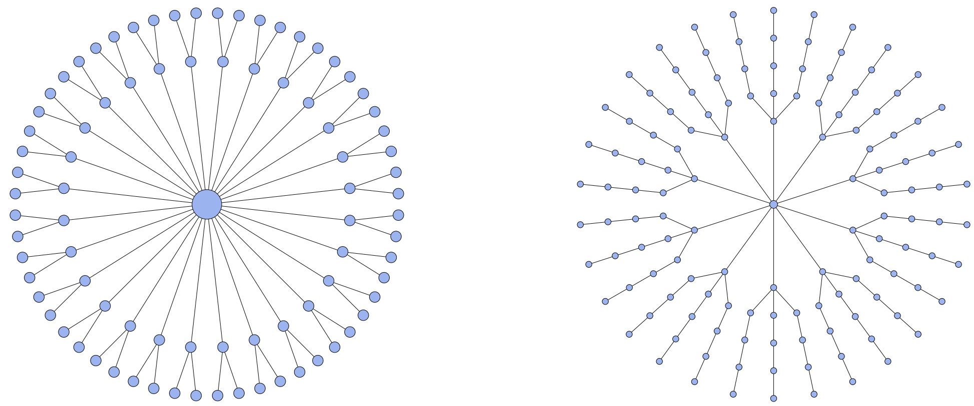
Let $v_i = b_i - t^*(b_i)$ and $k_i = k^*(b_i)$ where $t^*(s)$ and $k^*(s)$ are the minimizers under p^*. Let $\mathcal{I}$ be all continuous p such that $c'(0)s \leqslant p(s) \leqslant c(s)$ for all $s \in [0,1]$. Then we have the following result similar to Proposition 3.1.

Proposition 4.1. *If $c'(0) > 0$ and $g(k) \to \infty$ as $k \to \infty$, then (24) has a unique solution $p^* \in \mathcal{I}$ and $(p^*, \{v_i\}, \{k_i\})$ is an equilibrium for the production network.*

In the appendix, we show that the production network model also fits in our general dynamic programming framework developed in Section 2. In particular, assumptions A_1–A_5 hold so that Theorem 2.1 can be applied to the Bellman equation (24). Therefore, there exists a unique solution p^* that can be computed by value function iteration. We then prove that p^* induces an equilibrium allocation. Theorem 2.2 can also be used to show the monotonicity of p^*.

Figure 3 plots two production networks with different transaction costs, where each node corresponds to a firm in the network and the one in the center is the most downstream firm.[15] The size of each node is proportional to the size of the firm,

[15] We set $c(v) = v^{1.5}$ and $g(k) = 0.0001(k-1)^{1.5}$.



(A) $\tau = 0.2$ (B) $\tau = 0.05$

FIGURE 3. Examples of production networks.

represented by the sum of assembly and transaction costs. Figure 3 shows that more downstream firms are larger and have more upstream suppliers. Comparing panels (A) and (B), we can see that lower transaction costs increase the number of firms involved in the production network, encouraging the expansion of snakes. This is in line with the model prediction of Baldwin and Venables (2013) that decreasing frictions leads to a finer fragmentation of the production.

Tyazhelnikov's (2019) model of international production chains also shares some features with the model above. His model nests both snakes and spiders. Each firm makes optimal decision conditional on its production location at the next stage. If we interpret market transactions as offshoring, the multiple upstream supplier model becomes a model in which firms decide to produce parts of a production chain in any number of countries. Tyazhelnikov's (2019) prediction that firms choose to organize their production in large clusters to save on trade and unbundling costs is also present here, where firms complete production in-house to save on transactions costs. His prediction that trade liberalization leads to more fragmentation of the production process is reflected in the comparison between (A) and (B) in Figure 3.[16]

5. Application: Knowledge and Communication

Many firms are characterized by a pyramidal structure, in which employees are organized into management layers with each layer smaller than the previous one.

[16]The model developed in this section also bears some similarities with the endogenous production network model in Acemoglu and Azar (2020), where it is found that reducing distortions associated with an ad valorem tax leads to denser production networks. A reduction of transaction costs in our model captures the same idea, as shown in Figure 3.

These features have been modeled in the pioneering work by Garicano (2000) and the following literature. The key idea in Garicano's theory of hierarchical organization of knowledge is a trade-off between the cost of acquiring problem solving knowledge and the cost of communicating with others for help. His model highlights the impact that information and communication technology has on organizational design such as the number of management layers and the scope of production of workers in each layer. In this section, we solve a version of Garicano's model using the dynamic programming theory from Section 2.

Consider a model where production of a firm requires its employees solve a set of problems denoted by $[0, 1]$. Following Garicano (2000) (Section V.F), we suppose that there is a market for knowledge within the firm and each management layer solves a profit maximization problem. Suppose that employees at management layer i are assigned problems $m_i \in [0, 1]$. They learn to solve z_i at cost $c(z_i)$ and pass on the remainder $m_{i+1} = m_i - z_i$ to the next management layer $i + 1$ for help. This incurs additional communication costs τ that are proportional to the value of problems assigned to layer $i + 1$. Let $p\colon [0, 1] \to \mathbb{R}$ be the function of the value of problems in the internal market. Then, profits of the ith management layer are

$$\pi(m_i, z_i) = p(m_i) - (1 + \tau)p(m_i - z_i) - c(z_i),$$

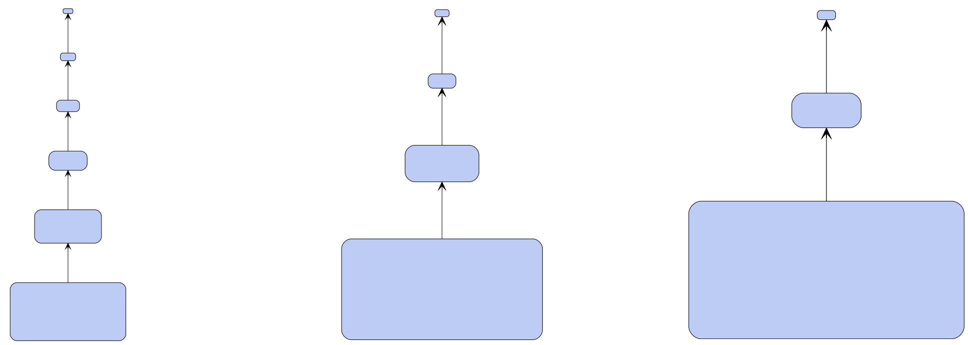
where $p(m_i)$ is the value of problems assigned to layer i, $(1 + \tau)p(m_i - z_i)$ is the cost of communicating and assigning unsolved problems to the next layer, and $c(z_i)$ is the cost of learning to solve z_i.

Setting profits to zero and minimizing with respect to m_{i+1} yield the equation

$$p(m_i) = \min_{m_{i+1} \leqslant m_i} \{c(m_i - m_{i+1}) + (1 + \tau)p(m_{i+1})\}.$$

This parallels the Bellman equation (4) in the negative discount dynamic programming in Section 2.1.2.

Let n be the number of employees in a given layer, and suppose that n is related to learning to solve z via $z = f(n)$. In other words, for a given range of problems z, the number of employees required to solve z is $n = f^{-1}(z)$. Assume that f is strictly increasing, strictly concave, and continuously differentiable with $f(0) = 0$, and that $c(z) = wn = wf^{-1}(z)$ for some wage rate w. Then the assumptions in Section 2.1.2 are satisfied if we let $\ell = c$ and $\beta = 1 + \tau$. The Euler equation (EU) implies that the optimal sequence $\{z_i\}$ is decreasing, so is the number of employees at each layer as $n_i = c(z_i)/w$. This replicates Garicano's result that the top management layer



(A) $\tau = 0.2$ (B) $\tau = 0.4$ (C) $\tau = 0.6$

FIGURE 4. Optimal organizational structures.

has the smallest number of employees and each layer below is larger than the one above.

The Euler equation (EU) suggests that the pyramidal structure of the span of control arises in equilibrium where each tier of management acquires knowledge up to the point where the marginal cost of learning to solve problems within the tier equals the marginal cost of communicating and assigning unsolved problems to the next layer. The envelope condition (EN) implies $p'(m_i) = c'(z_i)$, which says that, in equilibrium, the marginal value of problems assigned to a management layer equals the marginal cost of learning to solve problems within the tier.[17]

Figure 4 plots the optimal organizational structures of three firms given by the model above.[18] Each node corresponds to one management layer, who asks the layer above for help, and its size is proportional to the number of employees in that layer. As shown in the graphs, each firm has a pyramidal structure and higher communication costs increase the relative knowledge acquisition of lower layers and reduce the number of layers.

6. CONCLUSION

This paper shows how competitive equilibria in a range of production chain and network models can be recovered as the solution to a dynamic programming problem. Equilibrium prices are identified with the value function of a dynamic program, while competitive allocations of tasks across firms are identified with choices under

[17]This result is analogous to (17) for the production chain model and reminiscent of Coase's theory of the firm in the context of knowledge organization within a firm.

[18]We set $c(z) = z^{1.2}$ and $m_0 = 1$.

the optimal policy. Dynamic programming methods were then brought to bear on both the theory of the firm and the structure of production networks, providing new insights, as well as new analytical and computational methods. In addition to production problems, we also consider related competitive problems from economic geography and the problem of firm management. We hope the same ideas will prove useful in many other related problems.

7. Remaining Proofs

7.1. Proofs for Section 2.

Proof of Theorem 2.1. By A_1 and Berge's theorem of the maximum, Tw is continuous. Hence T maps $c\mathbb{R}^X$ to itself. It follows directly from A_2 that T is *isotone* on $c\mathbb{R}^X$, in the sense that $u \leqslant v$ implies $Tu \leqslant Tv$. Conditions A_4–A_5 and the isotonicity of T imply that, when $\varphi \leqslant w \leqslant \psi$, we have $\varphi \leqslant T\varphi \leqslant Tw \leqslant T\psi \leqslant \psi$. In particular, T is an isotone self-map on $\mathcal{I}$.

The Bellman operator is also concave on $\mathcal{I}$, in the sense that

$$0 \leqslant \lambda \leqslant 1 \text{ and } u, v \in \mathcal{I} \text{ implies } \lambda Tu + (1-\lambda)Tv \leqslant T(\lambda u + (1-\lambda)v). \tag{25}$$

Indeed, fixing such λ, u, v and applying A_3, we have

$$\min_{a \in G(x)} \{\lambda L(x, a, u) + (1-\lambda)L(x, a, v)\} \leqslant \min_{a \in G(x)} L(x, a, \lambda u + (1-\lambda)v)$$

for all $x \in X$. Since, for any pair of real valued functions f, g we have $\min_a f(a) + \min_a g(a) \leqslant \min_a(f(a) + g(a))$, it follows that (25) holds.

The preceding analysis shows that T is an isotone concave self-map on $\mathcal{I}$. In addition, by A_4 and A_5, we have $T\psi \leqslant \psi$ and $T\varphi \geqslant \varphi + \varepsilon(\psi - \varphi)$ for some $\varepsilon > 0$. Since $\mathcal{I}$ is an order interval in the positive cone of the Banach space $(c\mathbb{R}^X, \|\cdot\|)$, and since that cone is normal and solid, the first two claims in Theorem 2.1 are now confirmed via Theorem 2.1.2 of Zhang (2012). The final claim is due to Berge's theorem of the maximum. $\square$

Proof of Theorem 2.2. The first part of the theorem follows directly from the fact that $ic\mathbb{R}^X$ is a closed subspace. The proof is omitted. To prove the strict convexity of w^*, it suffices to show that Tw is strictly convex for all $w \in cc\mathbb{R}^X$ since $cc\mathbb{R}^X$ is a closed subspace of $c\mathbb{R}^X$. Pick any $x_1, x_2 \in X$ with $x_1 < x_2$ and any $\lambda \in (0, 1)$.

Let $x_\lambda = \lambda x_1 + (1-\lambda)x_2$. Pick any $w \in cc\mathbb{R}^X$ and let $\pi_w \colon X \to A$ be such that $(Tw)(x) = L(x, \pi_w(x), w)$. It follows that

$$\begin{aligned}
\lambda(Tw)(x_1) + (1-\lambda)(Tw)(x_2) &= \lambda L(x_1, \pi_w(x_1), w) + (1-\lambda)L(x_2, \pi_w(x_2), w) \\
&> L(x_\lambda, \lambda\pi_w(x_1) + (1-\lambda)\pi_w(x_2), w) \\
&\geqslant L(x_\lambda, \pi_w(x_\lambda), w) = (Tw)(x_\lambda),
\end{aligned}$$

where the first inequality holds because $(x, a) \mapsto L(x, a, w)$ is strictly convex and the second inequality holds because F_G is convex. Therefore, w^* is strictly convex. Strict convexity of L then implies that π^* is single-valued.

Since $\pi^*(x) \in \operatorname{int} G(x)$ and G is continuous, there exists an open neighborhood D of x such that $\pi^*(x) \in \operatorname{int} G(y)$ for all $y \in D$. Define $W(y) := L(y, \pi^*(x), w^*)$ for all $y \in D$. Then $W(y) \geqslant w^*(y)$ for all $y \in D$ and $W(x) = w^*(x)$. Since W is convex and differentiable on D, differentiability of w^* and (8) then follow from Benveniste and Scheinkman (1979). $\square$

We say that a dynamic programming problem has the *monotone increase* property if $-\infty < \varphi(x) \leqslant L(x, a, \varphi)$ for all $(x, a) \in F_G$ and Assumption 2.2 are satisfied. We state two useful lemmas from Bertsekas (2013).

Lemma 7.1 (Proposition 4.3.14, Bertsekas (2013))**.** *Let the monotone increase property hold and assume that the sets*

$$G_k(x, \lambda) := \{x \in G(x) \mid L(x, a, T^k\varphi) \leqslant \lambda\}$$

are compact for all $x \in X$, $\lambda \in \mathbb{R}$, and k greater than some integer $\bar{k}$. If $w \in \mathbb{R}^X_+$ satisfies $\varphi \leqslant w \leqslant \bar{w}$, then $\lim_{n\to\infty} T^n w = \bar{w}$. Furthermore, there exists an optimal stationary policy.

Lemma 7.2 (Proposition 4.3.9, Bertsekas (2013))**.** *Under the monotone increase property, a stationary policy π is optimal if and only if $T_\pi \bar{w} = T\bar{w}$.*

Proof of Theorem 2.3. Theorem 2.1 implies that $\lim_{n\to\infty} T^n \varphi = w^*$. To prove $w^* = \bar{w}$, it suffices to show that the conditions of Lemma 7.1 hold and $\varphi \leqslant \bar{w}$.

It follows from A_5 that $\varphi(x) \leqslant (T\varphi)(x) \leqslant L(x, a, \varphi)$ for all $(x, a) \in F_G$. Therefore, the monotone increase property is satisfied. Since T is a self-map on $c\mathbb{R}^X$, to check the conditions of Lemma 7.1, it suffices to prove that the set

$$G(x, \lambda) := \{x \in G(x) \mid L(x, a, w) \leqslant \lambda\}$$

is compact for any $w \in c\mathbb{R}^X$, $x \in X$, and $\lambda \in \mathbb{R}$. Since $a \mapsto L(x, a, w)$ is continuous by A_1, $L(x, \cdot, w)^{-1}((-\infty, \lambda])$ is a closed set. Since G is compact-valued, $G(x, \lambda)$ is compact. It remains to show that $\varphi \leqslant \bar{w}$. By A_5 and the definition of T, we have for any $\mu = (\pi_0, \pi_1, \dots) \in \mathcal{M}$, $\varphi \leqslant T^n \varphi \leqslant T_{\pi_0} T_{\pi_1} \dots T_{\pi_n} \varphi$ for all $n \in \mathbb{N}$. Then by definition, $\varphi \leqslant w_\mu$ for all $\mu \in \mathcal{M}$. Taking the infimum gives $\varphi \leqslant \bar{w}$ and there exists an stationary optimal policy. Lemma 7.1 then implies that $w^* = \bar{w}$. The principle of optimality follows directly from Lemma 7.2. □

7.2. **Proofs for the Negative Discount Dynamic Program.** Let $\mathcal{F}$ be the set of increasing convex functions in $\mathcal{I}$. Throughout the proofs, we regularly use the alternative expression for T given by

$$Tw(x) = \min_{0 \leqslant y \leqslant x} \{\ell(x-y) + \beta w(y)\} \tag{26}$$

Also, given $w \in \mathcal{F}$, define

$$\pi_w(x) = \operatorname*{arg\,min}_{0 \leqslant a \leqslant x} \{\ell(a) + \beta w(x-a)\}$$

and

$$\sigma_w(x) := \operatorname*{arg\,min}_{0 \leqslant y \leqslant x} \{\ell(x-y) + \beta w(y)\} = x - \pi_w(x). \tag{27}$$

These functions are clearly well-defined, unique and single-valued. Let $\sigma = \sigma_{w^*}$ and $\pi = \pi_{w^*}$. Let η be the constant defined by

$$\eta := \max\{0 \leqslant x \leqslant \hat{x} : \ell'(x) \leqslant \beta \ell'(0)\}. \tag{28}$$

We begin with several lemmas. The proof of the first lemma is trivial and hence omitted.

Lemma 7.3. *We have $\eta > 0$ if and only if $\ell'(0) > 0$. If $\eta < \hat{x}$, then $\ell'(\eta) = \beta \ell'(0)$.*

Lemma 7.4. *If $w \in \mathcal{F}$, then $\sigma_w(x) = 0$ if and only if $x \leqslant \eta$.*

Proof. First suppose that $x \leqslant \eta$. Seeking a contradiction, suppose there exists a $y \in (0, x]$ such that $\ell(x-y) + \beta w(y) < \ell(x)$. Since $w \in \mathcal{F}$ we have $w(y) \geqslant \ell'(0)y$ and hence

$$\beta w(y) \geqslant \beta \ell'(0) y \geqslant \ell'(\eta) y.$$

Since $x \leqslant \eta$, this implies that $\beta w(y) \geqslant \ell'(x)y$. Combining these inequalities gives $\ell(x-y) + \ell'(x)y < \ell(x)$, contradicting convexity of ℓ.

Now suppose that $\sigma_w(x) = 0$. We claim that $x \leqslant \eta$, or, equivalently $\ell'(x) \leqslant \beta \ell'(0)$. To prove $\ell'(x) \leqslant \beta \ell'(0)$, observe that since $w \in \mathcal{F}$ we have $w(y) \leqslant \ell(y)$, and hence

$$\ell(x) \leqslant \ell(x-y) + \beta w(y) \leqslant \ell(x-y) + \beta \ell(y) \quad \text{for all } y \leqslant x.$$

It follows that
$$\frac{\ell(x) - \ell(x-y)}{y} \leqslant \frac{\beta \ell(y)}{y} \quad \text{for all} \quad y \leqslant x.$$
Taking the limit gives $\ell'(x) \leqslant \beta \ell'(0)$. □

Proposition 7.5. *If $\ell'(0) > 0$, then the conclusions of Theorem 2.1 are valid for the Bellman operator T in (5).*

Proof of Proposition 7.5. Let $A = X = [0, \hat{x}]$, $G(x) = [0, x]$ and $L(x, a, w) = \ell(a) + \beta w(x - a)$. Conditions A_1–A_3 obviously hold. Condition A_4 holds since $\min_{0 \leqslant a \leqslant x}\{\ell(a) + \beta \ell(x - a)\} \leqslant \ell(x)$. For condition A_5, note that $L(x, a, \varphi) = \ell(a) + \beta \ell'(0)(x - a)$. Then $T\varphi = \ell$ if $x < \eta$ and $(T\varphi)(x) = \ell(\eta) + \beta \ell'(0)(x - \eta)$ if $x \geqslant \eta$. For $x < \eta$, $T\varphi - \varphi = \psi - \varphi$ so we can choose any $\varepsilon \leqslant 1$. For $x \geqslant \eta$,
$$\begin{aligned}(T\varphi)(x) - \varphi(x) &= \ell(\eta) + \beta \ell'(0)(x - \eta) - \ell'(0)x \\ &= \ell(\eta) - \ell'(0)\eta + (\beta - 1)\ell'(0)(x - \eta) \\ &\geqslant \ell(\eta) - \ell'(0)\eta = (\psi - \varphi)(\eta).\end{aligned}$$
Since $\psi - \varphi$ is increasing, we can choose any $\varepsilon \leqslant \bar{\varepsilon}$ where $(\psi - \varphi)(\eta) = \bar{\varepsilon}(\psi - \varphi)(\hat{x})$. The proposition thus follows from Theorem 2.1. □

Proposition 7.6. *The fixed point w^* of the negative discount Bellman operator T in (5) is strictly increasing, strictly convex, and continuously differentiable on $(0, \hat{x})$. The policy correspondence π^* is single-valued and satisfies $(w^*)'(x) = \ell'(\pi(x))$.*

Proof of Proposition 7.6. Consider the alternative expression for T in (26). Since ℓ is strictly convex, $(x, y) \mapsto \ell(x - y) + \beta w(y)$ is strictly convex for all $w \in cc\mathbb{R}^X$. Hence, part 1 of Assumption 2.1 holds. Evidently Tw is strictly convex for all $w \in \mathcal{F}$.

Next we show that Tw is strictly increasing for all $w \in \mathcal{F}$. Pick any $w \in \mathcal{F}$ and $x_1 \leqslant x_2$. For ease of notation, let $y_i = \sigma_w(x_i)$ for $i \in \{1, 2\}$. If $y_2 \leqslant x_1$, then
$$\begin{aligned}(Tw)(x_1) &= \ell(x_1 - y_1) + \beta w(y_1) \\ &\leqslant \ell(x_1 - y_2) + \beta w(y_2) \\ &< \ell(x_2 - y_2) + \beta w(y_2) = (Tw)(x_2)\end{aligned}$$
where the first inequality holds since y_2 is available when y_1 is chosen and the second inequality holds since ℓ is strictly increasing. If $y_2 > x_1$, we first consider the case of $x_1 + y_2 < x_2$. Then $(Tw)(x_2) > \ell(x_1) + \beta w(y_2) \geqslant \ell(x_1) \geqslant (Tw)(x_1)$. For the case

of $x_1 + y_2 \geqslant x_2$, we have $0 \leqslant y_1' \leqslant x_1 < y_2$ where $y_1' = x_1 + y_2 - x_2$. Since w is not constant, $w \in \mathcal{F}$ implies that w is strictly increasing. It follows that

$$\begin{aligned}(Tw)(x_1) &= \ell(x_1 - y_1) + \beta w(y_1) \\ &\leqslant \ell(x_1 - y_1') + \beta w(y_1') \\ &< \ell(x_2 - y_2) + \beta w(y_2) = (Tw)(x_2).\end{aligned}$$

Therefore, T is a self-map on $\mathcal{F}$ and Tw is strictly increasing and strictly convex for all $w \in \mathcal{F}$. Theorem 2.2 then implies that w^* is strictly increasing and strictly convex.

Since ℓ is differentiable, part 2 of Assumption 2.1 holds. Theorem 2.2 then implies that w^* is differentiable and $(f^*)'(x) = \ell'(x - \sigma(x))$ whenever $\sigma(x)$ is interior. Lemma 7.4 implies that $w^*(x) = \ell(x)$ and thus $(f^*)'(x) = \ell'(x)$ when $x \leqslant \eta$; when $x > \eta$, σ is interior and $(f^*)'(x) = \ell'(x - \sigma(x))$. Since σ is continuous, $(f^*)'$ is continuous. Therefore, w^* is continuously differentiable on $(0, \hat{x})$ and $(f^*)'(x) = \ell'(\pi(x))$. □

The next lemma further characterizes π and σ.

Lemma 7.7. *Let $w \in \mathcal{F}$. If x_1, x_2 satisfy $0 < x_1 \leqslant x_2$, then $\sigma_w(x_1) \leqslant \sigma_w(x_2)$ and $\pi_w(x_1) \leqslant \pi_w(x_2)$. Moreover, if $x \geqslant \eta$, then $\pi_w(x) \geqslant \eta$; if $x \leqslant \eta$, then $\pi_w(x) = x$.*

Proof. Pick any $w \in \mathcal{F}$. Since ℓ and w are convex, the maps $(x, a) \mapsto \ell(a) + \beta w(x - a)$ and $(x, y) \mapsto \ell(x - y) + \beta w(y)$ both satisfy the single crossing property. It follows from Theorem 4′ of Milgrom and Shannon (1994) that π_w and σ_w are increasing.

For the last claim, since π_w is increasing, Lemma 7.4 implies that, if $\eta \leqslant x$, then $\pi_w(x) \geqslant \pi_w(\eta) = \eta - \sigma_w(\eta) = \eta$; and if $x \leqslant \eta$, then $\pi_w(x) = x - \sigma_w(x) = x$. □

The following lemma characterizes the solution to (2) and is useful when showing the equivalence between (2) and (4).

Lemma 7.8. *If $\{a_t\}$ is a solution to (2), then $\{a_t\}$ is monotone decreasing and $a_{T+1} = 0$ if and only if $a_T \leqslant \eta$.*

Proof. The first claim is obvious, because if $\{a_t\}$ is a solution to (2) with $a_t < a_{t+1}$, then, given that $\beta > 1$, swapping the values of these two points in the sequence will preserve the constraint while strictly decreasing total loss. Regarding the second claim, since $\{a_t\}$ is monotone decreasing, it suffices to check the case $a_T > 0$. To this end, suppose to the contrary that $\{a_t\}$ is a solution to (2) with $0 < a_T < \eta$ and

$a_{T+1} > 0$. Consider an alternative feasible sequence $\{\hat{a}_t\}$ defined by $\hat{a}_T = a_T + \varepsilon$, $\hat{a}_{T+1} = a_{T+1} - \varepsilon$ and $\hat{a}_t = a_t$ for other t. If we compare the values of these two sequences we get

$$\begin{aligned}\sum_{t=0}^{\infty} \beta^t \ell(a_t) - \sum_{t=0}^{\infty} \beta^t \ell(\hat{a}_t) &= \beta^T [\ell(a_T) - \ell(a_T + \varepsilon)] + \beta^{T+1} [\ell(a_{T+1}) - \ell(a_{T+1} - \varepsilon)] \\ &= \varepsilon \beta^T \left\{ -\frac{\ell(a_T + \varepsilon) - \ell(a_T)}{\varepsilon} + \beta \frac{\ell(a_{T+1} - \varepsilon) - \ell(a_{T+1})}{-\varepsilon} \right\}.\end{aligned}$$

The term inside the parenthesis converges to

$$-\ell'(a_T) + \beta \ell'(a_{T+1}) > -\ell'(\eta) + \beta \ell'(0) \geqslant 0,$$

where the first inequality follows from $a_T \leqslant \eta$, $a_{T+1} > 0$ and strict convexity of ℓ; and the second inequality is by the definition of η. We conclude that for ε sufficiently small, the difference $\sum_{t=0}^{\infty} \beta^t \ell(a_t) - \sum_{t=0}^{\infty} \beta^t \ell(\hat{a}_t)$ is positive, contradicting optimality.

Finally we check the claim $a_{T+1} = 0 \implies a_T \leqslant \eta$. Note that if $\eta = \hat{x}$ then there is nothing to prove, so we can and do take $\eta < \hat{x}$. Seeking a contradiction, suppose instead that $a_{T+1} = 0$ and $a_T > \eta$. Consider an alternative feasible sequence $\{\hat{a}_t\}$ defined by $\hat{a}_T = a_T - \varepsilon$, $\hat{a}_{T+1} = \varepsilon$ and $\hat{a}_t = a_t$ for other t. In this case we have

$$\sum_{t=0}^{\infty} \beta^t \ell(a_t) - \sum_{t=0}^{\infty} \beta^t \ell(\hat{a}_t) = \varepsilon \beta^T \left\{ \frac{\ell(a_T - \varepsilon) - \ell(a_T)}{-\varepsilon} - \beta \frac{\ell(\varepsilon) - \ell(0)}{\varepsilon} \right\}.$$

The term inside the parentheses converges to

$$\ell'(a_T) - \beta \ell'(0) > \ell'(\eta) - \beta \ell'(0) = 0,$$

where the final equality is due to $\eta < \hat{x}$ and Lemma 7.3. Once again we conclude that for ε sufficiently small, the difference $\sum_{t=0}^{\infty} \beta^t \ell(a_t) - \sum_{t=0}^{\infty} \beta^t \ell(\hat{a}_t)$ is positive, contradicting optimality. □

Proposition 7.9. *For the negative discount dynamic program, the sequence $\{a_t^*\}$ defined by $x_0 = \hat{x}$, $x_{t+1} = x_t - \pi^*(x_t)$ and $a_t^* = \pi^*(x_t)$ is the unique solution to (2). Moreover, $W = w^*$.*

Proof of Proposition 7.9. To show the equivalence between (2) and (4), we first show that (2) is equivalent to $\bar{w} = \inf_{\mu \in \mathcal{M}} w_\mu$ where w_μ is as defined in (11). Suppose that the optimal policy is $\mu = (\pi_0, \pi_1, \dots)$ and we let $\sigma_t(x) = x - \pi_t(x)$. Then we have

$$\begin{aligned}\bar{w}(\hat{x}) = w_\mu(\hat{x}) = \ell[\pi_0(\hat{x})] + \beta \ell[\pi_1 \sigma_0(\hat{x})] + \beta^2 \ell[\pi_2 \sigma_1 \sigma_0(\hat{x})] + \dots \\ + \limsup_{t \to \infty} \beta^k \ell'(0) \sigma_{t-1} \sigma_{t-2} \cdots \sigma_0(\hat{x}). \quad (29)\end{aligned}$$

It is clear that $\bar{w}$ is finite. Therefore, the optimal policy must satisfy $\sigma_t \to 0$, otherwise the last term in (29) would go to infinity. Let $a_t = \pi_t \sigma_{t-1} \dots \sigma_0(\hat{x})$. We claim that $\{a_t\}$ solves (2). Suppose not and the solution to (2) is $\{a'_t\}$. Then by Lemma 7.8, $a'_t = 0$ for all $t > T$ for some T. Thus we can construct a policy μ' that reproduces $\{a'_t\}$ and gives a lower loss. This is a contradiction. Conversely, suppose that the solution to (2) is $\{a_t\}$. Using the same argument, we can show that the policy that gives rise to $\{a_t\}$ is an optimal policy. Therefore, $W = \bar{w}$.

Next we show that $w^* = \bar{w}$ using Theorem 2.3. That Assumption 2.2 holds was shown in Section 2.4.1. It follows from Theorem 2.3 that $w^* = \bar{w}$, there exists an stationary optimal policy, and the Bellman's principle of optimality holds. Since π^* satisfies $T_{\pi^*} w^* = T w^*$, π^* is a stationary optimal policy.

Theorems 2.1 and 2.2 implies that π^* is continuous and single-valued. It then follows from the principle of optimality that $\{a^*_t\}$ is the unique solution to (2). $\square$

Proposition 7.10. *For all $n \in \mathbb{N}$ and increasing convex $w \in \mathcal{I}$, we have*

$$T^n w(x) = w^*(x) \quad \textit{whenever} \quad x \leqslant n\eta.$$

Proposition 7.10 implies uniform convergence in *finite* time. In particular, for $n \geqslant \hat{x}/\eta$ we have $T^n w = w^*$ everywhere on $[0, \hat{x}]$. Note that this bound $\hat{x}/\eta$ is independent of the initial condition w.

Proof of Proposition 7.10. It suffices to show that if $f, g \in \mathcal{F}$, then $T^k f = T^k g$ on $[0, k\eta]$. We prove this by induction.

To see that $T^1 f = T^1 g$ on $[0, \eta]$, pick any $x \in [0, \eta]$ and recall from Lemma 7.4 that if $h \in \mathcal{F}$ and $x \leqslant \eta$, then $Th(x) = \ell(x)$. Applying this result to both f and g gives $Tf(x) = Tg(x) = \ell(x)$. Hence $T^1 f = T^1 g$ on $[0, \eta]$ as claimed.

Turning to the induction step, suppose now that $T^k f = T^k g$ on $[0, k\eta]$, and pick any $x \in [0, (k+1)\eta]$. Let $h \in \mathcal{F}$ be arbitrary, let π_h be the h-greedy function, and let $\sigma_h(x) := x - \pi_h(x)$. By Lemma 7.7, we have $\pi_h(x) \geqslant \eta$, and hence

$$\sigma_h(x) \leqslant x - \eta \leqslant (k+1)\eta - \eta \leqslant k\eta.$$

In other words, given function h, the optimal choice at x is less than $k\eta$. Since this is true for both $h = T^k f$ and $h = T^k g$, we have

$$T^{k+1} f(x) = \min_{0 \leqslant y \leqslant x} \{\ell(x-y) + \beta T^k f(y)\} = \min_{0 \leqslant y \leqslant k\eta} \{\ell(x-y) + \beta T^k f(y)\}.$$

Using the induction step we can now write

$$T^{k+1}f(x) = \min_{0\leqslant y\leqslant k\eta}\{\ell(x-y)+\beta T^k g(y)\} = \min_{0\leqslant y\leqslant x}\{\ell(x-y)+\beta T^k g(y)\}.$$

The last expression is just $T^{k+1}g(x)$, and we have now shown that $T^{k+1}f = T^{k+1}g$ on $[0,(k+1)\eta]$. The proof is complete. $\square$

Proof of Proposition 2.4. Since $\ell'(0)=0$, (EU) is equivalent to $\beta\ell'(a^*_{t+1}) = \ell'(a^*_t)$.

Sufficiency. Let $x^*_0 = \hat{x}$ and $x^*_t = x^*_{t-1} - a^*_{t-1}$ for $t \geqslant 1$. Let $\{a_t\}$ be any feasible sequence. Let $x_0 = \hat{x}$ and $x_t = x_{t-1} - a_{t-1}$. It suffices to prove that

$$D := \lim_{T\to\infty}\sum_{t=0}^{T}\beta^t[\ell(a^*_t)-\ell(a_t)] \leqslant 0.$$

Since ℓ is convex, we have

$$\begin{aligned} D &= \lim_{T\to\infty}\sum_{t=0}^{T}\beta^t[\ell(x^*_t - x^*_{t+1}) - \ell(x_t - x_{t+1})] \\ &\leqslant \lim_{T\to\infty}\sum_{t=0}^{T}\beta^t\ell'(a^*_t)(x^*_t - x_t - x^*_{t+1} + x_{t+1}). \end{aligned}$$

Since $x_0 = x^*_0$, rearranging gives

$$D \leqslant \lim_{T\to\infty}\sum_{t=0}^{T}\beta^t(x^*_{t+1}-x_{t+1})[\beta\ell'(a^*_{t+1}) - \ell'(a^*_t)] - \beta^T\ell'(a^*_T)(x^*_{T+1}-x_{T+1}).$$

Since $\beta\ell'(a^*_{t+1}) = \ell'(a^*_t)$, the summation is zero and $\beta^T\ell'(a^*_T) = \ell'(a^*_0)$. We have

$$D \leqslant -\lim_{T\to\infty}\ell'(a^*_0)(x^*_{T+1}-x_{T+1}).$$

Since $\{a_t\}$ and $\{a^*_t\}$ are feasible, x_{T+1} and x^*_{T+1} go to zero when $T\to\infty$. Therefore, $D \leqslant 0$.

Existence and Uniqueness. Since $\{a^*_t\}$ is feasible and satisfies $\beta\ell'(a^*_{t+1}) = \ell'(a^*_t)$ for all t, we have

$$\hat{x} = \sum_{t=0}^{\infty}a^*_t = \sum_{t=0}^{\infty}(\ell')^{-1}\left(\frac{1}{\beta^t}\ell'(a^*_0)\right) =: g(a^*_0),$$

where $(\ell')^{-1}$ is well defined on $[0, \lim_{x\to\infty}\ell'(x)]$ because ℓ is increasing, strictly convex, and $\ell'(0)=0$. Hence, g is well defined on $\mathbb{R}_+$ and $g(a^*_0)$ is continuous and strictly increasing in a^*_0. Since $g(0)=0$ and $g(\hat{x}) > \hat{x}$, there exists a unique $a^*_0 > 0$ such that $\{a^*_t\}$ satisfying $\beta\ell'(a^*_{t+1}) = \ell'(a^*_t)$ is feasible, $a^*_t > 0$ for all t, and $\{a^*_t\}$ is strictly decreasing. That $\{a^*_t\}$ is an optimal solution then follows from the sufficiency part. Since ℓ is strictly convex, the solution is unique.

Necessity. Since we have pinned down a unique solution of (2) which satisfies $\beta\ell'(a^*_{t+1}) = \ell'(a^*_t)$, the condition is also necessary. $\square$

7.3. **Proofs for Section 3.**

Proof of Proposition 3.1. We must verify that $(W, \{a^*_i\})$ satisfies Definition 3.1. We first consider the case of $\ell'(0) > 0$. By Propositions 7.5 and 7.9, the value function W is a solution to the Bellman equation (4), and hence satisfies

$$W(s) = \min_{0 \leqslant v \leqslant s} \{c(v) + (1+\tau)W(s-v)\} \quad \text{for all } s \in [0,1]. \tag{30}$$

By Proposition 7.6, it lies in the class $\mathcal{F}$ of increasing, convex and continuous functions $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $c'(0)s \leqslant f(s) \leqslant c(s)$ for all $s \in \mathbb{R}_+$. In addition, with $\{x_i\}$ as the optimal state process (see Proposition 7.9), we have,

$$W(x_i) = \{c(a^*_i) + (1+\tau)W(x_{i+1})\} \quad \text{for all } i \geqslant 0. \tag{31}$$

We need to show that 1–3 of Definition 3.1 hold when $p = W$ and $v_i = a^*_i$ for all $i \in \mathbb{Z}$. Part 1 is immediate because $W \in \mathcal{F}$ and all functions in $\mathcal{F}$ must have this property, while Part 2 follows directly from (30). To see that Part 3 of Definition 3.1 also holds, let $b_i = x_i$. By the definition of the state process, the sequence $\{b_i\}$ then corresponds to the downstream boundaries of a set of firms obeying task allocation $\{a^*_i\}$. The profits of firm i are $\pi_i = W(b_i) - c(a^*_i) - (1+\tau)W(b_{i+1})$. By (31) and $b_i = x_i$, we have $\pi_i = 0$ for all i. Hence Part 3 of Definition 3.1 also holds, as was to be shown.

If $\ell'(0) = 0$, part 1 follows from the definition of the value function (14). By Proposition 2.4, for any t with $0 \leqslant t \leqslant 1$, there exists a unique optimal allocation $\{a^*_{t,j}\}$ such that $W(t) = \sum_j \beta^j \ell(a^*_{t,j})$, and $\sum_j a^*_{t,j} = t$. Since $\{s-t, a^*_{t,0}, a^*_{t,1}, \ldots\}$ is a feasible allocation at stage s with $t \leqslant s \leqslant 1$, part 2 follows from the definition of the value function. To see part 3, let $b_0 = 1$ and $b_i = b_{i-1} - a^*_{i-1}$. By Proposition 2.4, we have $\ell'(a^*_i) = (1+\tau)\ell'(a^*_{i+1})$. Since $\sum_{i=j}^{\infty} a^*_i = b_j$ for all j, it follows again from Proposition 2.4 that $\{a^*_i\}_{i=j}^{\infty}$ is an optimal allocation for stage b_j. Therefore, $p(b_i) = \sum_{j=0}^{\infty}(1+\tau)^j c(a^*_{i+j}) = c(a^*_i) + (1+\tau)p(b_{i+1})$ for all i. Hence, $\pi_i = 0$ for all i. $\square$

7.4. **Proofs for Section 4.**

Proof of Proposition 4.1. To study this problem in the framework of Theorem 2.1, we set $X = [0, \hat{x}]$, $A = [0, \hat{x}] \times \mathbb{N}$, $G(x) = [0, x] \times \mathbb{N}$, and

$$L(x, a, w) = c(x-t) + g(k) + (1+\tau)kp(t/k) \qquad a = (t, k).$$

Since $g(k) \to \infty$ as $k \to \infty$, we can restrict $G(x)$ to be $[0, x] \times \{1, 2, \ldots, \bar{k}\}$ so that G is compact-valued. Under the conditions of Proposition 4.1, it can be shown that A_1–A_5 hold with $\psi = c$ and $\varphi(s) = c'(0)s$ (see Yu and Zhang (2019)). Then, Theorem 2.1 implies that the Bellman equation (24) has a unique solution p^* in $\mathcal{I}$, $T^n p \to p^*$ for all $p \in \mathcal{I}$ where

$$(Tp)(s) := \min_{\substack{0 \leqslant t \leqslant s \\ k \in \mathbb{N}}} L(x, a, w),$$

and t^* and k^* exist. We need only verify that $(p^*, \{v_i\}, \{k_i\})$ given by $v_i = b_i - t^*(b_i)$, $k_i = k^*(b_i)$ and $b_{i+1} = (b_i - v_i)/k_i$ is an equilibrium, the definition of which is given in Section 4.2.

Since $p^* \in \mathcal{I}$, $p(0) = 0$. Since p^* satisfies (24), part (ii) of the definition is also satisfied. To see that part (iii) holds, note that

$$\begin{aligned} p^*(b_i) &= c(b_i - t^*(b_i)) + g(k^*(b_i)) + (1+\tau)k^*(b_i)p^*\left(\frac{t^*(b_i)}{k^*(b_i)}\right) \\ &= c(v_i) + g(k_i) + (1+\tau)k_i p^*\left(\frac{b_i - v_i}{k_i}\right). \end{aligned}$$

It follows that $\pi_i = 0$ for all $i \in \mathbb{Z}$ where π_i is as defined in (23). This completes the proof. □

References

Acemoglu, D. and P. D. Azar (2020): "Endogenous production networks," *Econometrica*, 88, 33–82.

Antràs, P. and A. De Gortari (2020): "On the geography of global value chains," *Econometrica*.

Baldwin, R. and A. J. Venables (2013): "Spiders and snakes: Offshoring and agglomeration in the global economy," *Journal of International Economics*, 90, 245–254.

Becker, G. S. and K. M. Murphy (1992): "The division of labor, coordination costs, and knowledge," *The Quarterly Journal of Economics*, 107, 1137–1160.

Bellman, R. (1957): *Dynamic programming*, Academic Press.

Benveniste, L. M. and J. A. Scheinkman (1979): "On the differentiability of the value function in dynamic models of economics," *Econometrica*, 47, 727–732.

Bertsekas, D. P. (2013): *Abstract dynamic programming*, Athena Scientific Belmont, MA.

——— (2017): *Dynamic programming and optimal control*, vol. 4, Athena Scientific.

Bloise, G. and Y. Vailakis (2018): "Convex dynamic programming with (bounded) recursive utility," *Journal of Economic Theory*, 173, 118–141.

Boehm, J. and E. Oberfield (2018): "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," Tech. rep., National Bureau of Economic Research.

Buera, F. J. and Y. Shin (2011): "Self-insurance vs. self-financing: A welfare analysis of the persistence of shocks," *Journal of Economic Theory*, 146, 845–862.

Carvalho, V. M. and B. Grassi (2019): "Large firm dynamics and the business cycle," *American Economic Review*, 109, 1375–1425.

Christaller, W. (1933): *Central Places in Southern Germany*, Translation into English by Carlisle W. Baskin in 1966, Englewood Cliffs, NJ: Prentice-Hall.

Christiano, L., M. Eichenbaum, and S. Rebelo (2011): "When is the government spending multiplier large?" *Journal of Political Economy*, 119, 78–121.

Clementi, G. L. and B. Palazzo (2016): "Entry, exit, firm dynamics, and aggregate fluctuations," *American Economic Journal: Macroeconomics*, 8, 1–41.

Coase, R. H. (1937): "The nature of the firm," *Economica*, 4, 386–405.

Coe, N. M. and H. W.-C. Yeung (2015): *Global production networks: Theorizing economic development in an interconnected world*, Oxford University Press.

Costinot, A., J. Vogel, and S. Wang (2013): "An elementary theory of global supply chains," *The Review of Economic Studies*, 80, 109–144.

Epstein, L. and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–69.

Fally, T. and R. Hillberry (2018): "A Coasian model of international production chains," *Journal of International Economics*, 114, 299–315.

Farlow, A. (2020): "Beyond COVID-19: Supply Chain Resilience Holds Key to Recovery," Tech. rep., Baker McKenzie.

Gabaix, X. (2009): "Power laws in economics and finance," *Annu. Rev. Econ.*, 1, 255–294.

Gabaix, X. and Y. M. Ioannides (2004): "The evolution of city size distributions," in *Handbook of regional and urban economics*, Elsevier, vol. 4, 2341–2378.

Garicano, L. (2000): "Hierarchies and the organization of knowledge in production," *Journal of Political Economy*, 108, 874–904.

Garicano, L. and E. Rossi-Hansberg (2006): "Organization and inequality in a knowledge economy," *The Quarterly Journal of Economics*, 121, 1383–1435.

Hills, T. S., T. Nakata, and S. Schmidt (2019): "Effective lower bound risk," *European Economic Review*, 120, 103321.

Hsu, W.-T. (2012): "Central place theory and city size distribution," *The Economic Journal*, 122, 903–932.

Hsu, W.-T., T. J. Holmes, and F. Morgan (2014): "Optimal city hierarchy: A dynamic programming approach to central place theory," *Journal of Economic Theory*, 154, 245–273.

Kamihigashi, T., K. Reffett, and M. Yao (2015): "An application of Kleene's fixed point theorem to dynamic programming," *International Journal of Economic Theory*, 11, 429–434.

Kikuchi, T., K. Nishimura, and J. Stachurski (2018): "Span of control, transaction costs, and the structure of production chains," *Theoretical Economics*, 13, 729–760.

Kocherlakota, N. R. (1990): "On the 'discount' factor in growth economies," *Journal of Monetary Economics*, 25, 43–47.

Kremer, M. (1993): "The O-ring theory of economic development," *The Quarterly Journal of Economics*, 108, 551–575.

Levine, D. K. (2012): "Production chains," *Review of Economic Dynamics*, 15, 271–282.

Loewenstein, G. and D. Prelec (1991): "Negative time preference," *The American Economic Review*, 81, 347–352.

Loewenstein, G. and N. Sicherman (1991): "Do workers prefer increasing wage profiles?" *Journal of Labor Economics*, 9, 67–84.

Loewenstein, G. and R. H. Thaler (1989): "Anomalies: intertemporal choice," *Journal of Economic Perspectives*, 3, 181–193.

Lucas, R. E. (1978): "On the size distribution of business firms," *The Bell Journal of Economics*, 508–523.

Martins-da Rocha, F. V. and Y. Vailakis (2010): "Existence and uniqueness of a fixed point for local contractions," *Econometrica*, 78, 1127–1141.

Mehra, R. and E. C. Prescott (1985): "The equity premium: A puzzle," *Journal of monetary Economics*, 15, 145–161.

Milgrom, P. and C. Shannon (1994): "Monotone comparative statics," *Econometrica: Journal of the Econometric Society*, 157–180.

Rincón-Zapatero, J. P. and C. Rodríguez-Palmero (2003): "Existence and uniqueness of solutions to the Bellman equation in the unbounded case," *Econometrica*, 71, 1519–1555.

Stokey, N. and R. E. Lucas (1989): *Recursive Methods in Economic Dynamics (with EC Prescott)*, Harvard University Press.

Thaler, R. (1981): "Some empirical evidence on dynamic inconsistency," *Economics Letters*, 8, 201–207.

Tyazhelnikov, V. (2019): "Production Clustering and Offshoring," Tech. rep., University of Sydney.

Varian, H. R. (2002): "If There Was a New Economy, Why Wasn't There a New Economics?" *New York Times*, 17, C2.

Yu, M. and J. Zhang (2019): "Equilibrium in production chains with multiple upstream partners," *Journal of Mathematical Economics*, 83, 1–10.

Zhang, Z. (2012): *Variational, topological, and partial order methods with their applications*, vol. 29, Springer Science & Business Media.