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

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

May 27, 2021

Abstract. We show that competitive equilibria in a range of models related to production networks can be recovered as solutions to dynamic programs. Although these programs fail to be contractive, we prove that they are tractable. As an illustration, we treat Coase's theory of the firm, equilibria in production chains with transaction costs, and equilibria in production networks with multiple partners. We then show how the same techniques 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
JEL Classification: C61, D21, D90

1. Introduction

Production networks have grown rapidly in size and complexity, in line with advances in communications, supply chain management and transportation technology (see, e.g., Coe and Yeung (2015)). These large and complex networks are sensitive to uncertainty, trade disputes, transaction costs and other frictions. Firms routinely shift production and task allocation across networks, in order to mitigate risk or

[1]We gratefully acknowledge valuable comments from our editors and referees, as well as Pol Antràs, Yongsung Chang, Davin Chor, Ryo Jinnai, Young Sik Kim, Yuta Takahashi and Makoto Saito. We also acknowledge financial support from Korea University (K1922081), BrainKorea21 Plus (K1327408 and T192201), JSPS Grant-in-Aid for Scientific Research (20H05633 and 16H03598), Singapore Ministry of Education Academic Research Fund R-122-000-140-112 and Australian Research Council Award DP120100321. Email contacts: tomookikuchi@waseda.jp, nishimura@rieb.kobe-u.ac.jp, john.stachurski@anu.edu.au, zhangjunnan1224@gmail.com

[a] Graduate School of Asia-Pacific Studies, Waseda University; [b] Research Institute of Economics and Business Administration, Kobe University, and RIETI; [c] Research School of Economics, Australian National University; [d] Center for Macroeconomic Research at School of Economics, and Wang Yanan Institute for Studies in Economics, Xiamen University

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.

Dynamic programming provides one methodology for analyzing such equilibria. While dynamic programming is typically used to study *dynamic* models (see, e.g., Stokey and Lucas (1989)), it can also be applied to static models by reinterpreting the time parameter as an index over firms or other decision making entities, as seen in, for example, Garicano and Rossi-Hansberg (2006), Hsu et al. (2014), Tyazhelnikov (2019), and Antràs and De Gortari (2020). 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.

This research agenda faces a technical hurdle: the dynamic programs most naturally mapped to the competitive allocation problems we wish to consider usually fail to be 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. In this paper, we circumvent the need for contractivity by drawing on dynamic programming methods originally developed to solve recursive preference problems.[2]

The contributions of this paper fall into two parts. The first is providing a theory of dynamic programming in a loss-minimization setting where discount rates are negative. 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 application, we build on an analytical framework for analyzing allocation of tasks across firms first developed by Coase (1937). Subsequently, Kikuchi et al. (2018), Fally and Hillberry (2018) and Yu and Zhang (2019) developed Coasian models in which firms trade off coordination costs within the firm against transaction

[2]See, for example, Epstein and Zin (1989), Bloise and Vailakis (2018) or Marinacci and Montrucchio (2019). 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.

costs outside the firm. We show that competitive equilibria in these models can be recovered as solutions to dynamic programs and use the associated envelope condition to provide insight on some of the foundational conjectures of Coase (1937).

In the remainder of the paper, we then apply similar methods to study a range of additional applications, including settings where Coasian transaction costs are replaced by failures in production or costly transportation, as found, for example, in Levine (2012) and Costinot et al. (2013); models of knowledge organization and optimal management structures originally due to Garicano (2000); the analysis of central place theory in Hsu et al. (2014); and the configuration of general (nonsequential) production networks in the spirit of Baldwin and Venables (2013), Kikuchi et al. (2018), Yu and Zhang (2019) and Tyazhelnikov (2019).

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 negative discount dynamic programming theory developed here provides a unifying methodology and brings tools to bear on understanding the structure of the networks where firms, cities and managers coordinate production.

Regarding our technical contribution, the closest existing work in the economic literature is Bloise and Vailakis (2018), who treat noncontractive dynamic programming problems that arise from recursive utility. 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 apply a fixed point result of Du (1989) to 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]

The remainder of this paper is structured as follows. In Section 2, we study a dynamic optimization problem under negative discounting 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 show that our model can also be used to understand organization of knowledge

[3]This 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.

within a firm. In Section 5 we extend our model to expand the scope of applications to more complex networks. Section 6 concludes. Most proofs are deferred to the appendix.

2. Negative Discount Dynamic Programming

In this section, we study a dynamic optimization problem in which an agent minimizes a flow of losses under negative discounting. While our main aim is to develop techniques for calculating equilibria in production networks, the topic of negative discount loss minimization does have some independent value.[4]

Consider an agent who takes action a_t in period t with loss $\ell(a_t)$. We interpret a_t as effort and $\ell(a_t)$ as disutility. Her optimization problem is, for some $\hat{x} > 0$,

$$\min_{\{a_t\}} \sum_{t=0}^{\infty} \beta^t \ell(a_t) \quad \text{s.t. } a_t \geqslant 0 \text{ for all } t \geqslant 0 \text{ and } \sum_{t=0}^{\infty} a_t = \hat{x}. \tag{1}$$

Throughout this section, we suppose that

$$\beta > 1, \ \ell(0) = 0, \ \ell' > 0 \ \text{ and } \ell'' > 0. \tag{2}$$

The convexity in ℓ encourages the agent to defer some effort. Negative discounting ($\beta > 1$) has the opposite effect. We call problem (1) under the assumptions in (2) a *negative discount dynamic program.*[5]

[4] For example, Thaler (1981), Loewenstein and Prelec (1991) and Loewenstein and Sicherman (1991) document separate instances of such phenomena. 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).

[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 disutility, 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.

2.1. **A Recursive View.** We can express the problem recursively by introducing a state process $\{x_t\}$ that starts at $\hat{x}$ and tracks the amount of tasks remaining. Set $x_{t+1} = x_t - a_t$ and $x_0 = \hat{x}$. The Bellman equation for this problem is

$$w(x) = \inf_{0 \leqslant a \leqslant x} \{\ell(a) + \beta w(x-a)\}. \tag{3}$$

The Bellman operator is

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

The Bellman operator is not a supremum norm contraction because $\beta > 1$.[6] Nevertheless, we can show that T is well behaved, with a unique fixed point, after we restrict its domain to a suitable candidate class $\mathcal{I}$. To this end, we set

$$X := [0, \hat{x}], \quad \varphi(x) := \ell'(0)x \quad \text{and} \quad \psi(x) := \ell(x).$$

Let $\mathcal{I}$ be all continuous $w\colon X \to \mathbb{R}$ with $\varphi \leqslant w \leqslant \psi$. These upper and lower bounds have natural interpretations. Since completing all remaining tasks at once is in the choice set, its value $\ell(x)$ is an upper bound of the minimized value. Regarding the lower bound $\ell'(0)x$, this is the value that could be obtained if $\beta = 1$ (no discounting) and the agent, having no time constraint, subdivided without limit.

Proposition 2.1. *The Bellman operator has a unique fixed point w^* in $\mathcal{I}$ and $T^n w \to w^*$ as $n \to \infty$ for all $w \in \mathcal{I}$. Moreover,*

1. *w^* is strictly increasing, strictly convex, and continuously differentiable, and*
2. *The policy $\pi^*(x) := \arg\min_{0 \leqslant a \leqslant x}\{\ell(a) + \beta w^*(x-a)\}$ is single-valued and satisfies*

$$(w^*)'(x) = \ell'(\pi^*(x)) \qquad (0 < x < \hat{x}). \tag{5}$$

Stability of the fixed point is derived from the monotonicity and concavity of the Bellman operator in Appendix A.3. In Proposition A.11 we show that the convergence $T^n w \to w^*$ always converges in finite time.

2.2. **Equivalence.** So far, we have solved the Bellman equation (3) and derived properties of its solutions. However, it is not clear whether the Bellman equation can characterize the solution to the dynamic optimization problem (1), since the constraint $\sum_t a_t = \hat{x}$ is not in the Bellman equation. We turn to this issue now.

[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 $+\infty$.

Let

$$W(x) := \min \left\{ \sum_{t=0}^{\infty} \beta^t \ell(a_t) \;\; : \;\; \{a_t\} \in \mathbb{R}_+^\infty \text{ and } \sum_{t=0}^{\infty} a_t = x \right\} \tag{6}$$

be the *value function* of the optimization problem (1). The next proposition shows that $W = w^*$, the fixed point of T, and that the policy correspondence π^* solves (1). The proof can be found in Appendix A.3.

Proposition 2.2. *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 (1). Moreover, $W = w^*$.*

The envelope condition (5) now evaluates to

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

for all $t \geqslant 0$, 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 $t \geqslant 0$, 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.

Equation (EU) implies that $\{a_t^*\}$ is a decreasing sequence. This agrees with our intuition, since future losses are given greater weight than current losses.

2.3. **Additional Results.** Instead of assuming $\ell' > 0$ as in (2), we can treat the case $\ell'(0) = 0$, which has hitherto been excluded:

Proposition 2.3. *When $\ell'(0) = 0$, a feasible sequence $\{a_t^*\}$ solves (1) if and only if (EU) holds. This sequence is unique, decreasing, and satisfies $a_t^* > 0$ for all t.*

Proposition 2.3 shows that the Euler equation (EU) becomes necessary and sufficient for optimality when $\ell'(0) = 0$. In fact, (EU) can be reduced to $\beta \ell'(a_{t+1}^*) = \ell'(a_t^*)$ in this case, which helps derive analytical solutions for some of the applications.

As the above results suggest, the set of tasks will be completed in finite time if and only if $\ell'(0) > 0$. The proof is in the appendix.

[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).

3. Application: Production Chains

Now we turn to applications of our negative discount dynamic program motivated by production problems. We begin with linear production chains.

3.1. A Coasian Production Chain.

In this section we consider a version of the Coasian models developed by Kikuchi et al. (2018), Fally and Hillberry (2018) and Yu and Zhang (2019). We show how competitive equilibrium in these models can be calculated using the dynamic programming theory from Section 2.

3.1.1. *Set Up.*

Consider a market with many price-taking firms, each of which is either inactive or part of the production of a single good. To produce a unit of this good requires implementing a unit mass of tasks. The cost for any one firm of implementing an interval of length $v \in [0,1]$ is denoted $c(v)$, where c is increasing, strictly convex, continuously differentiable, and satisfies $c(0) = 0$.[8]

Firms face transaction costs, as a wedge between cost to the buyer and payment received by the seller.[9] Transaction costs fall on buyers, so that, for a transaction with face value f, the seller receives f and the buyer pays $(1+\tau)f$ with $\tau > 0$.[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 $i \geqslant 0$. Then, profits of the ith firm are

$$\pi_i = p(b_i) - c(v_i) - (1+\tau)p(b_{i+1}). \tag{7}$$

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 (7). The pair $(p, \{v_i\})$ is called an *equilibrium* for the production chain if

1. $p(0) = 0$,

---

[8] Unlike Kikuchi et al. (2018), we allow $c'(0) = 0$.

[9] This follows Kikuchi et al. (2018) and also studies such as Boehm and Oberfield (2020), 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.

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

Conditions 1–3 eliminate profits for active firms and prevent entry by inactive firms.

3.1.2. *Solution by Dynamic Programming.* 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 (3). Moreover, iterating on this relation yields the price of the final good

$$p(1) = \sum_{i \geqslant 0} (1+\tau)^i c(v_i), \tag{8}$$

which is analogous to the total loss in (1). These facts lead us to a version of the negative discount dynamic program introduced in Section 2 where a (fictitious) agent seeks to minimize $\sum_{i \geqslant 0}(1+\tau)^i c(a_i)$ subject to $\sum_{i \geqslant 0} a_i = 1$. By construction, any feasible action path is also a feasible allocation of tasks in the production chain.

Since $\ell = c$ and $\beta = 1 + \tau$, the assumptions in (2) are satisfied. Hence there exists a unique solution $\{a_i^*\}$ by Proposition 2.2. Let W be the corresponding value function given by (6). The next proposition shows that the solution to this dynamic program is precisely the competitive equilibrium of the Coasian production chain.

Proposition 3.1. *If $p = W$ and $v_i = a_i^*$ for all $i \geqslant 0$, then $(p, \{v_i\})$ is an equilibrium for the production chain.*

For firm with downstream boundary b_i, the envelope condition (EN) yields

$$p'(b_i) = c'(v_i). \tag{9}$$

Since v_i is the optimal range of tasks implemented in-house by firm i in equilibrium, this expresses 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. The Euler equation (EU) also implies that $\{v_i\}$ is decreasing, so 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 $f(k, n) \geqslant v$. 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 (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 2.3 applies and (EU) yields $a^*_{i+1} = \theta a^*_i$ for all $i \geqslant 0$, 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 (6) gives the price function

$$p(x) = W(x) = \kappa\,(1-\theta)^{(1-\eta)/\eta}\,x^{1/\eta}. \tag{10}$$

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

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 (10) becomes linear.[11]

3.2. **Specialization and Failure Probabilities.** Production processes typically involve a series of complementary tasks. Mistakes in any one task 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). These papers show how equilibrium allocations can serve to mitigate the potentially exponential cost of failures in long production chains.[12] In this section, we show that these ideas are also amenable to analysis using the negative discount dynamic program from Section 2.

Consider, as before, a competitive market where producers implement a mass of tasks contained in $[0,1]$. We drop the assumption of positive transaction costs and

[11]The above result on the size of firms is related to Antràs and De Gortari (2020), who prove it is optimal to locate relatively downstream stages of production in relatively central locations where trade costs are lower. 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, due to (EU).

[12]For example, 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.

replace it with positive probability of defects.[13] 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 facing price function p are

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

This parallels the profit function (7) from the Coasian production chain model. If we adopt the Cobb–Douglass technology from Section 3.1.3, then the price of the final good is

$$p^*(1) = \kappa \left(1 - (1+\tau)^{\eta/(\eta-1)}\right)^{(1-\eta)/\eta}. \tag{11}$$

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

$$\hat{p}^*(1) = \kappa \sum_{i=0}^{N} (1+\tau)^i \left(\frac{1}{N}\right)^{1/\eta} = \kappa \frac{(1+\tau)^N - 1}{(1+\tau) - 1} \left(\frac{1}{N}\right)^{1/\eta} = O((1+\tau)^N). \tag{12}$$

Now a small increase in τ 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 of prices with and without producers adjusting for failure probabilities.[14]

Thus, returning to the original model, we see that equilibrium prices induce producers to adjust to changes in failure probabilities, which optimally mitigates the potentially exponential impact of failures on the cost of the final good.

---

[13] Defects can alternatively be understood as iceberg costs, where some percentage of goods are lost in transporting them from one producer to the next.

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

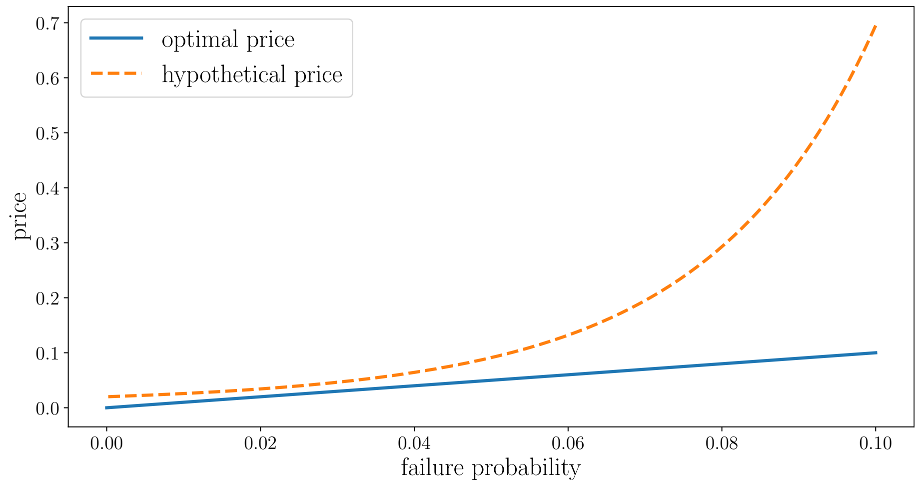


Figure 1. Final good price and failure probabilities.

4. Application: Knowledge and Communication

Many firms are characterized by a pyramidal structure, in which employees form management layers with each layer smaller than the previous one. These features were modeled by Garicano (2000), where hierarchical organization of knowledge involves a trade-off between the cost of acquiring problem solving knowledge and the cost of communicating with others for help. In this section, we solve a version of Garicano's model using the dynamic programming theory from Section 2.

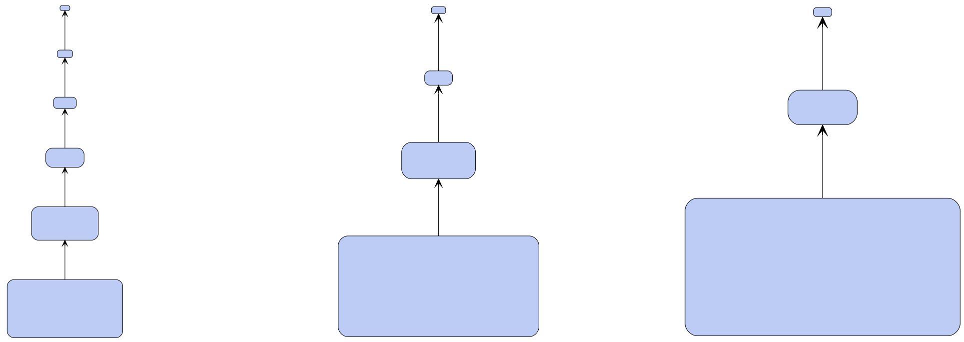
Consider a firm where production requires solving a set of problems. Employees at management layer i are assigned problems $m_i \in [0, 1]$. They learn to solve z_i at cost $c(z_i)$ and pass on the remainder $m_{i+1} = m_i - z_i$ to the next management layer $i+1$. This incurs additional communication costs that are proportional to the value of problems assigned to layer $i+1$ with coefficient τ.

Let $p\colon [0, 1] \to \mathbb{R}$ be a (fictitious) price function that assigns value to problems. 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} yields

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



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

FIGURE 2. Optimal organizational structures.

This parallels the Bellman equation (3) of the negative discount dynamic program in Section 2.

Suppose that n employees can learn to solve $z = f(n)$ problems. 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 (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, as is the number of employees in each layer as $n_i = c(z_i)/w$. This replicates Garicano's result that the top management layer has the smallest number of employees and each layer below is larger than the one above.

The Euler equation (EU) adds additional insight: each layer of management acquires knowledge up to the point where the marginal cost of learning 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 layer.[15]

Figure 2 plots the optimal organizational structures of three firms given by the model above.[16] 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

[15]This result is analogous to (9) for the production chain model and reminiscent of Coase's theory of the firm in the context of knowledge organization within a firm.

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

communication costs increase the relative knowledge acquisition of lower layers and reduce the number of layers.

5. Extension: Nonlinear Networks

In this section we treat more general network models that cannot be directly handled by the theory in Section 2. Unlike the linear chains discussed above, agents can interact with multiple partners. In Section 5.1, we study a problem from economic geography. In Section 5.2 we study chains with multiple upstream partners using a general dynamic programming theory developed in Appendix A.1.

5.1. Spatial Networks.

The distribution of city sizes shows remarkable regularity, as described by the rank-size rule.[17] One early attempt to match the empirical city size distribution is found in the central place theory of Christaller (1933). Hsu (2012) and Hsu et al. (2014) formalize Christaller's theory. In this section, we develop a model with similar insights by extending our earlier dynamic programming results.

Consider a government that opens competition for many developers to build cities to host a continuum of dwellers indexed by $[0, 1]$. Each developer can build a large city that hosts everyone or build a smaller city and pay 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 city building starts with a single developer, who is assigned the whole population, and ends with a network of cities consisting of multiple layers.

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

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

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

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

where every dweller is accommodated and every developer makes zero profits. The equilibrium price function satisfies

$$p(s) = \min_{0 \leqslant t \leqslant s} \left\{ c(s-t) + (1+\tau)kp\left(\frac{t}{k}\right) \right\}. \tag{13}$$

To find the equilibrium price function, we first solve a negative discount dynamic program and then show that its value function is the solution to (13).

Consider a dynamic optimization problem with value function given by

$$W(x) := \min_{\{v_i\}} \left\{ \sum_{i=0}^{\infty} (1+\tau)^i k^i c(v_i) : \{v_i\} \in \mathbb{R}_+^{\infty} \text{ and } \sum_{i=0}^{\infty} k^i v_i = x \right\}. \tag{14}$$

The problem in (14) is a modified version of (1) that also features negative discounting and a convex loss function. In the context of our city network model, (14) describes a social planner who minimizes the total cost of hosting the whole population, where v_i stands for the size of cities on layer i.

In what follows we let $c(s) = s^\gamma$ with $\gamma > 1$. To emulate the bifurcation process in Hsu (2012) and Hsu et al. (2014), we let $k = 2$. A similar argument to the proof of Proposition 2.3 gives the Euler equation

$$c'(v_i) = (1+\tau)c'(v_{i+1}). \tag{15}$$

Using this equation, it can be shown with some algebra that $v_i = \theta^i(1-2\theta)$ for $\theta := (1+\tau)^{1/(1-\gamma)} < 1/2$ and the value function is $W(s) = (1-2\theta)^{\gamma-1}s^\gamma$. It is straightforward to verify that $p = W$ satisfies (13). Hence, the value function for the social planner is also the equilibrium price function under which all developers make zero profits.

The Euler equation (15) 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 3 illustrates the optimal city hierarchy by placing cities according to Hsu (2012) and Hsu et al. (2014).[18] It replicates 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

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

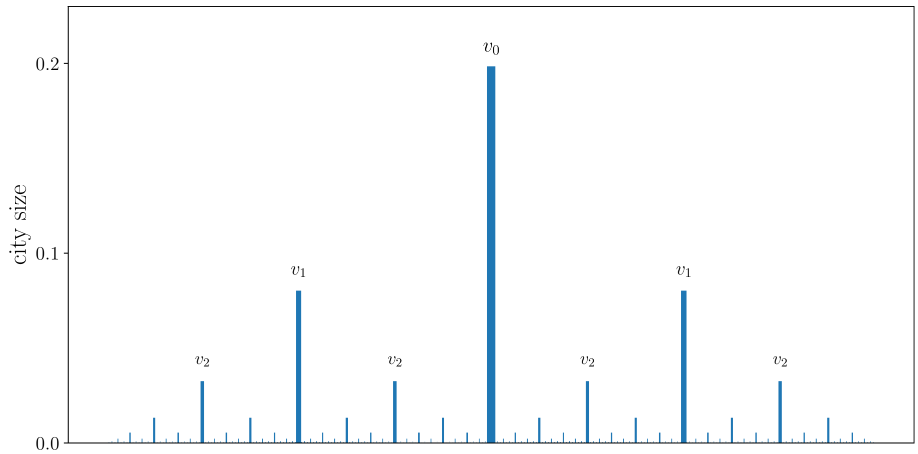


FIGURE 3. Illustration of optimal city hierarchy.

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.

5.2. **Snakes and Spiders.** Modern production networks are characterized by processes that are both sequential and non-sequential. Baldwin and Venables (2013) refer to the sequential processes as “snakes” and the non-sequential processes as “spiders”, and analyze how the location of different parts of a production chain is determined by unbundling costs of production across borders. Here we study a model of production networks featuring both snakes and spiders.

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. 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 integers $i \geqslant 0$ 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 $b_{i+1} = (b_i - v_i)/k_i$. We call $(p, \{v_i\}, \{k_i\})$ 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 $i \geqslant 0$ where

$$\pi_i := p(b_i) - c(v_i) - g(k_i) - (1+\tau)k_i p\left(\frac{b_i - v_i}{k_i}\right). \tag{16}$$

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

$$p(s) = \min_{\substack{0 \leqslant t \leqslant s \\ k \in \mathbb{N}}} \left\{c(s-t) + g(k) + (1+\tau)kp(t/k)\right\}. \tag{17}$$

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 $s \in [0,1]$.

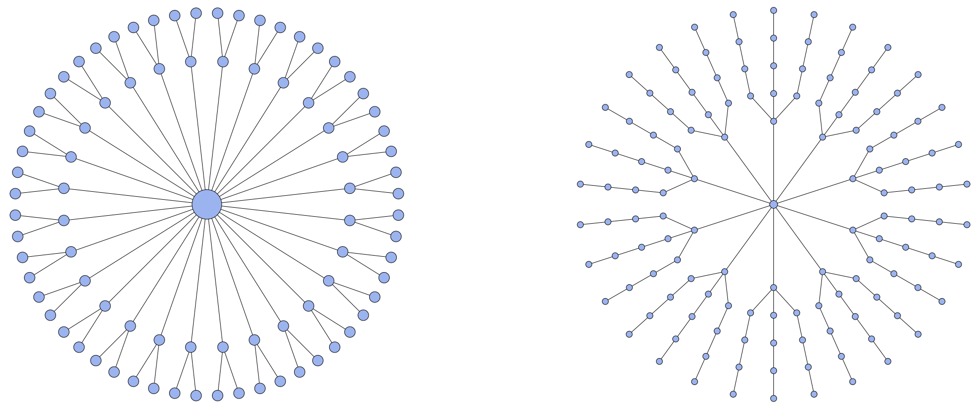
Proposition 5.1. *If $c'(0) > 0$ and $g(k) \to \infty$ as $k \to \infty$, then (17) has a unique solution $p^* \in \mathcal{I}$ and $(p^*, \{v_i\}, \{k_i\})$ is an equilibrium for the production network.*

In Appendix A.5, we show that the unique solution p^* can be computed by value function iteration. We then prove that p^* induces an equilibrium allocation. Theorem A.2 can also be used to show the monotonicity of p^*.

Figure 4 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.[19] The size of each node is proportional to the size of the firm, represented by the sum of assembly and transaction costs. Figure 4 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.[20]

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

[20]Tyazhelnikov's (2019) model of international production chains 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.



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

Figure 4. Examples of production networks.

6. Conclusion

This paper shows how competitive equilibria in a range of production chain and network models can be recovered as solutions to dynamic programming problems. Equilibrium prices are identified with the value function of a dynamic program, while competitive allocations across firms are identified with choices under the optimal policy. Dynamic programming methods are 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 firm management.

Apart from the model of snakes and spiders in Section 5.2, all of the problems faced by individual firms are convex. This assumption allowed us to obtain sharp results and useful characterizations. An important remaining task is to extend our results to a range of cases that feature non-convexities. This work is left for future research.

Appendix A. Appendix

A.1. **A General Dynamic Programming Framework.** In this section, we provide a general dynamic programming framework suitable for analyzing equilibria in production networks.

A.1.1. *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 $\|f\| := \sup_{x \in E} |f(x)|$.

Let X be a compact metric space. Let A be a metric space and let G be a nonempty, continuous, compact-valued correspondence from X to A. We understand $G(x)$ as the set of available actions $a \in A$ for an agent in 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, mapping $F_G \times \mathbb{R}^X$ 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 as a *dynamic program*.

The Bellman operator associated with such a pair is the operator T defined by

$$(Tw)(x) = \inf_{a \in G(x)} L(x, a, w) \qquad (w \in \mathbb{R}^X,\ x \in X). \tag{18}$$

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

A.1.2. *Fixed Point Results.* Fix a 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
$$\lambda L(x, a, u) + (1 - \lambda) L(x, a, v) \leqslant L(x, a, \lambda u + (1 - \lambda) v).$$
- 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 concavity. 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 Bellman operator (18) is well behaved under A_1–A_5 after restricting its domain to a suitable class of candidate solutions. To this end, let

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

Theorem A.1. *Let (L, G) be a dynamic program and let T be the Bellman operator defined in (18). If (L, G) satisfies A_1–A_5, then*

1. *T has a unique fixed point w^* in $\mathfrak{I}$.*
2. *For each $w \in \mathfrak{I}$, there exists an $\alpha < 1$ and $M < \infty$ such that*
$$\|T^n w - w^*\| \leqslant \alpha^n M \quad \textit{for all } n \in \mathbb{N}. \tag{19}$$
3. *$\pi^*(x) := \arg\min_{a \in G(x)} L(x, a, w^*)$ is upper hemicontinuous on X.*

The fixed point results in Theorem A.1 rely on the monotonicity and concavity of the Bellman operator. See Section A.2 for details of the arguments and the proof of the theorem.

Theorem A.1 does not discuss Bellman's principle of optimality. That task is left until Section A.1.4. 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.

A.1.3. *Shape and Smoothness Properties.* We now give conditions under which the solution to the Bellman equation associated with a 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. The following assumption is needed for convexity and differentiability.

Assumption A.1. In addition to A_1–A_5, the 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 A.2. *If Tw is strictly increasing for all $w \in ic\mathbb{R}^X$, then w^* is strictly increasing. If Assumption A.1 holds, then w^* is strictly convex, π^* is single-valued, w^* is differentiable on* $\operatorname{int} X$ *and*

$$(w^*)'(x) = L_x(x, \pi^*(x), w^*) \tag{20}$$

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

A.1.4. *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

$$(T_\pi w)(x) = L(x, \pi(x), w). \tag{21}$$

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

$$w_\mu(x) := \limsup_{n\to\infty} (T_{\pi_0} T_{\pi_1} \dots T_{\pi_n} \varphi)(x), \tag{22}$$

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

$$\bar{w}(x) := \inf_{\mu \in \mathcal{M}} w_\mu(x) \tag{23}$$

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

Assumption A.2. In addition to A_1–A_5, the 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 $w \geqslant \varphi$,

$$L(x, a, w + r) \leqslant L(x, a, w) + \beta r. \tag{24}$$

Part 1 of Assumption A.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 A.2 is analogous to the Blackwell's condition, with the significant exception that β in (24) is not restricted to be less than one.

Theorem A.3. *If Assumption A.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 A.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 A.1 is an optimal stationary policy.

A.2. **Proofs for the General Theory.** To prove Theorem A.1, we first give a fixed point theorem for monotone concave operators on a partially ordered Banach space due to Du (1989).[21]

Theorem A.4 (Du, 1989)**.** *Let P be a normal cone on a real Banach space E.*[22] *Suppose $u_0, v_0 \in E$ with $u_0 < v_0$ and $A : [u_0, v_0] \to E$ is an increasing concave operator. If $Au_0 \geqslant u_0 + \varepsilon(v_0 - u_0)$ for some $\varepsilon \in (0,1)$ and $Av_0 \leqslant v_0$, then A has a unique fixed point x^* in $[u_0, v_0]$. Furthermore, for any $x \in [u_0, v_0]$ and $n \in \mathbb{N}$, $\|A^n x - x^*\| \leqslant M(1-\varepsilon)^n$ for some $M > 0$.*

Proof of Theorem A.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 A.1 are now confirmed via Theorem A.4. The final claim is due to Berge's theorem of the maximum. $\square$

Proof of Theorem A.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 $\lambda \in (0,1)$.

[21] The theory of monotone concave operators dates back to Krasnosel'skii (1964). Similar treatments include, for example, Guo and Lakshmikantham (1988), Guo et al. (2004), and Zhang (2013).

[22] A cone $P \subset E$ is said to be normal if there exists $\delta > 0$ such that $\|x + y\| \geqslant \delta$ for all $x, y \in P$ and $\|x\| = \|y\| = 1$.

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 (20) then follow from Benveniste and Scheinkman (1979). $\square$

We say that a dynamic programming problem has the *monotone increasing* property if $-\infty < \varphi(x) \leqslant L(x,a,\varphi)$ for all $(x,a) \in F_G$ and Assumption A.2 are satisfied. We state two useful lemmas from Bertsekas (2013).

Lemma A.5 (Proposition 4.3.14, Bertsekas (2013)). *Let the monotone increasing 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 A.6 (Proposition 4.3.9, Bertsekas (2013)). *Under the monotone increasing property, a stationary policy π is optimal if and only if $T_\pi \bar{w} = T\bar{w}$.*

Proof of Theorem A.3. Theorem A.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 A.5 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 increasing property is satisfied. Since T is a self-map on $c\mathbb{R}^X$, to check the conditions of Lemma A.5, 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_2 and the monotone increasing property, we have for any $\mu = (\pi_0, \pi_1, \ldots) \in \mathcal{M}$, $\varphi \leqslant T_{\pi_0} T_{\pi_1} \ldots T_{\pi_n} \varphi$ for all $n \in \mathbb{N}$. Then by definition, $\varphi \leqslant w_\mu$ for all $\mu \in \mathcal{M}$. Taking the infimum gives $\varphi \leqslant \bar{w}$. Lemma A.5 then implies that $w^* = \bar{w}$ and there exists an optimal stationary policy. The principle of optimality follows directly from Lemma A.6. □

A.3. **Proofs for Section 2.** 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 A.7. *We have $\eta > 0$ if and only if $\ell'(0) > 0$. If $\eta < \hat{x}$, then $\ell'(\eta) = \beta \ell'(0)$.*

Lemma A.8. *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 $\ell'(x) \leqslant \beta\ell'(0)$. □

Proof of Proposition 2.1. 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 in Section A.1.2 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 first part of the proposition thus follows from Theorem A.1.

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 A.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 A.2 then implies that w^* is strictly increasing and strictly convex.

Since ℓ is differentiable, part 2 of Assumption A.1 holds. Theorem A.2 then implies that w^* is differentiable and $(w^*)'(x) = \ell'(x - \sigma(x))$ whenever $\sigma(x)$ is interior. Lemma A.8 implies that $w^*(x) = \ell(x)$ and thus $(w^*)'(x) = \ell'(x)$ when $x \leqslant \eta$; when $x > \eta$, σ is interior and $(w^*)'(x) = \ell'(x - \sigma(x))$. Since σ is continuous, $(w^*)'$ is continuous. Therefore, w^* is continuously differentiable on $(0, \hat{x})$ and $(w^*)'(x) = \ell'(\pi(x))$. □

The next lemma further characterizes π and σ.

Lemma A.9. *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 A.8 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 (1) and is useful when showing the equivalence between (1) and (3).

Lemma A.10. *If $\{a_t\}$ is a solution to (1), 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 (1) 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 (1) 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

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

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 A.7. 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. □

Proof of Proposition 2.2. To show the equivalence between (1) and (3), we first show that (1) is equivalent to $\bar{w} = \inf_{\mu \in \mathcal{M}} w_\mu$ where w_μ is as defined in (22). 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

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

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 (1). Suppose not and the solution to (1) is $\{a'_t\}$. Then by Lemma A.10, $a'_t = 0$ 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 (1) 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 A.3. Both conditions in Assumption A.2 can be verified for (L, G). Part 1 of Assumption A.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 A.3 applies. It follows from Theorem A.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 A.1 and A.2 imply that π^* is continuous and single-valued. It then follows from the principle of optimality that $\{a_t^*\}$ is the unique solution to (1). ☐

Proposition A.11. *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 A.11 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 A.11. 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 A.8 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 A.9, 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. ☐

Proof of Proposition 2.3. 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

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

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 as $T \to \infty$. Hence $D \leqslant 0$.

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 (1) which satisfies $\beta \ell'(a_{t+1}^*) = \ell'(a_t^*)$, the condition is also necessary. □

A.4. **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 2.1 and 2.2, the value function W is a solution to the Bellman equation (3), and hence satisfies

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

and W 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 2.2), 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 $i \geqslant 0$. 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 (6). By Proposition 2.3, 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 $\{s-t, a_{t,0}^*, a_{t,1}^*, \dots\}$ 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 2.3, 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 2.3 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. □

A.5. **Proofs for Section 5.**

Proof of Proposition 5.1. To study this problem in the framework of Section A.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 $[0, x] \times \{1, 2, \dots, \bar{k}\}$ so that G is compact-valued. Under the conditions of Proposition 5.1, 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 A.1 implies that the Bellman equation (17) 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 5.2.

Since $p^* \in \mathcal{I}$, $p(0) = 0$. Since p^* satisfies (17), 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 (16). This completes the proof. $\square$

References

Antràs, P. and A. De Gortari (2020): "On the geography of global value chains," *Econometrica*, 88, 1553–1598.

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

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.

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 (2020): "Misallocation in the market for inputs: enforcement and the organization of production," *The Quarterly Journal of Economics*, 135, 2007–2058.

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

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.

Du, Y. (1989): "Fixed points of a class of non-compact operators and applications," *Acta Mathematica Sinica*, 32, 618–627.

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.

Guo, D., Y. J. Cho, and J. Zhu (2004): *Partial Ordering Methods in Nonlinear Problems*, Nova Publishers.

Guo, D. and V. Lakshmikantham (1988): *Nonlinear Problems in Abstract Cones*, Academic Press.

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.

Krasnosel'skii (1964): *Positive Solutions of Operator Equations*, Noordhoff.

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.

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

Marinacci, M. and L. Montrucchio (2019): "Unique tarski fixed points," *Mathematics of Operations Research*, 44, 1174–1191.

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

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.

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

Zhang, Z. (2013): *Variational, Topological, and Partial Order Methods with Their Applications*, vol. 29 of *Developments in Mathematics*, Springer Berlin Heidelberg.