Payment-Chain Crises*

Saki Bigio[†] **Esteban Méndez**[‡] **Diana Van Patten**[§]

First Released Draft: December 2021

Abstract

This paper introduces an endogenous network of payment-chains into a business cycle model. Motivated by evidence of linked payments across firms in Costa Rica, we develop a framework where production orders form bilateral relations: some payments are executed immediately, while others—chained payments—are delayed until upstream payments are received. These chains capture real-world situations in which firms must wait to be paid before paying their own suppliers, leaving resources temporarily idle even when demand and capacity exist. In equilibrium, agents choose the amount of chained payments given interest rates and access to internal funds or credit lines. This choice determines the payment-chain network and aggregate total-factor productivity (TFP). The paper characterizes equilibrium dynamics and pecuniary externalities when agents internalize their own payment delays but not the delays imposed on others.

Keywords: Payments, Networks, Business Cycles

JEL Classification: E32, E42, G01

---

*We thank Fernando Alvarez, Luigi Bocola, Ricardo Lagos, Jennifer La'O, Ezra Oberfield, Michael Peters, Guillaume Rocheteau, Gary Richardson, Pierre-Olivier Weill, and Yuliy Sannikov as well as seminar participants at UC Irvine, UCLA, Princeton, SARG, NBER IFM and MM, CEMLA, and CESifo, for early discussions. Sebastian Merkel, Gadi Barlevy, and Romain Ranciere provided thoughtful discussions. We are indebted to Ken Miyahara and Luis Yepez for their outstanding assistance, multiple discussions, and their creativity in tackling proofs. The views expressed herein are those of the authors and do not necessarily represent the views of the Central Bank of Costa Rica (BCCR). All results have been reviewed by the BCCR to ensure no confidential information is disclosed.

[†] UCLA and NBER. Email: sbigio@econ.ucla.edu. Address: 315 Portola Plaza Los Angeles, CA 90095.

[‡] Central Bank of Costa Rica. Email: mendezce@bccr.fi.cr. Address: Avenida Central y Primera, Calles 2 y 4, San José 10101.

[§] Yale University and NBER. Email: diana.vanpatten@yale.edu. Address: 165 Whitney Avenue, New Haven, CT 06511.

1 Introduction

A classic narrative has it that during financial crises, productive resources often remain idle, not because of technical constraints or insufficient of demand, but because agents must wait for others to pay, who in turn are waiting on others. A simple example illustrates the idea. A general contractor is hired to build a commercial property. The contractor relies on a network of subcontractors—electricians, plumbers, and carpenters—who agree to progress with their work only after receiving partial payments. The contractor also places orders with material suppliers, who customize materials but are unwilling to ship them until payment completion or, at minimum, a proof of funds is provided. However, the contractor is waiting for a payment from a developer, who is waiting for a credit line approval. During crises, banks may withdraw their credit lines, disrupting the developer's funding sources, and therefore, the entire chain: the contractor cannot pay the electrician or the supplier; the electrician postpones work; the supplier holds off on shipping; and the project stalls. The slowdown is not due to a lack of labor, materials, or technical capacity, but rather to each party waiting to be paid by someone else. The cascading leave productive human and physical resources idle. National accounts record the outcome as a total factor productivity (TFP) decline.

Despite how natural such payment delays seem, they are rarely featured in macroeconomic models. Production is typically centralized with spot market exchanges and payment frictions abstracted away. While credit constraints appear often, the usual assumption is binary: agents are either constrained or unconstrained. What is missing is a tractable framework where economic activity stalls simply because agents are waiting for transactions to clear.

This paper develops a theory of payment-chain disruptions that fits that narrative. We motivate our theory by analyzing detailed microdata on all interbank business-to-bussiness payments in Costa Rica, along with the universe of transaction-level sales receipts. We document that over half of firms engage mainly in chained payments, payments whose execution requires the execution of earlier payments. Exploiting a natural experiment where bank accounts were suddenly frozen for some firms, we show that payment disruptions lead to real effects on output, and that these effects depend critically on whether the affected firms participated in payment chains.

Motivated by this evidence, we introduce a payment-chain production network that explicitly links transaction timing to production timing. In this framework, some orders are paid upfront (spot), while others, chained-orders, remain unfulfilled until preceding payments clear. As the fraction of chained orders rises, production delays cascade through longer chains, reducing measured TFP according to a tractable formula linking micro-level transaction assignments to macro-level outcomes.

The payment-chain production network is portable as it can be aggregated into seemingly conventional budget constraints. These aggregation reveals two potential sources of inefficiencies. One relates to the organization of payments themselves and the other through pecuniary externalities that occur when agents do not internalize how their expenditures affect the speed of transactions in the network.

To demonstrate the use of our theory, we embed this network into a business-cycle model with entrepreneurs and creditors. When entrepreneurs face limited short-term credit access, they place chained orders to but inputs, internalizing the effects of delays on to them, but not internalizing how their delays affect others. The model generates three regions: efficient steady states with minimal delays, temporary payment-chain crises that resolve as debt falls, and permanent crises with debt overhang where deleveraging incentives are insufficient.

The framework yields novel policy insights. A Ramsey planner recognizing payment externalities finds both that during payment-chain crises creditors spend too little on spot orders and whereas entrepreneurs spend too much on chained orders during crises. Fiscal policy can have positive multipliers during payments-chain crises—but only if the government makes spot expenditures. If government spends without making upfront payments, their expenditures are detrimental. The mechanism works, not through aggregate demand stimulus, but by speeding up payments, a payments reinterpretation of fiscal multipliers.

Literature Review. The paper connects with theories that underscore the sharp declines in aggregate TFP during crises. The link between financial crisis and TFP is not at all obvious because financial crises can manifest themselves through distortions labor-market distortions, not productivity. One branch of the literature explains declines in aggregate TFP through increased capital misallocation or capital utilization—see Meza and Quintin (2007), Pratap and Urrutia (2012) or Oberfield (2013). A common finding is that these models can only partially explain TFP declines once disciplined with data. Our paper offers a different mechanism where financial conditions impact TFP through a slowdown of payments and economic activity.

The paper also falls at the crossroads of several areas, the monetary-payments literature, the economic-networks literature, and the literature on aggregate-demand externalities. The issue of how payment instruments affect production is a classic theme.[1] Recent work focuses on how the distribution funds, and not the quantity of funds, affects production—see Lippi et al. (2015), Rocheteau et al. (2016), and Brunnermeier and Sannikov (2017). Like this literature, the distribution of funding affects allocations, and the main distinction is that we focus on delays in sequential payments.

Sequential payments appear in many other studies. The payments-chain network is inspired by Townsend (1980) or Kiyotaki and Moore (1997a), which can be thought of models where production is distorted within a vertical supply chain.[2] Recent work by Hardy et al. (2022) and Bocola (2022) contrasts external funding against trade credit along the supply chain.[3] Relative to these papers, there are two distinctions: First, we consider a notion of timing where orders are unfulfilled until earlier payments are executed. Second, the network of payments emerges as an outcome of endogenous to expenditure decisions.

With respect to the networks literature, the paper connects with models with endogenous network formation. The contribution relative to that literature is modest, as network formation is not strategic. By contrast, in Oberfield (2018), a network is formed through strategic partnerships; in Kopytov et al. (2022) and Elliott et al. (2022) firms form strategic links anticipating possible supply-chain breakdowns. Here, the network is randomly formed, but the distribution of chains is endogenous to financial decisions. As in Elliott et al. (2014), Alvarez and Barlevy (2021) and Taschereau-Dumouchel (2022), there are externalities here too. In those models, externalities occur when individual defaults provoke subsequent defaults. Here, externalities occur through payment delays.

Finally, the paper connects with models of aggregate demand externalities. An early

[1]Lucas and Stokey (1987) analyzes a stochastic cash-in-advance economy; Kiyotaki and Wright (1989) studies trade with indivisible tokens; Lagos and Wright (2005) a model with divisible money and explicit trading arrangements. See also Shi (1997); Lagos et al. (2011); Lagos and Rocheteau (2009); Li et al. (2012); Nosal and Rocheteau (2011); Rocheteau (2011) for many other directions in that area.

[2]Other models of sequential payments include sequences of payments with spatial separation, Freeman (1996a) and Green (1999), which study sequential transactions in overlapping generation environments, La'O (2015), which studies a circular flow of transactions, or Guerrieri and Lorenzoni (2009), which studies sequential transactions in a Lagos and Wright-type environment. Bigio and La'O (2013) considers the propagation of financial shocks that induce misallocation in a production network.

[3]See also Biais and Gollier (1997) for an earlier model of trade credit. The importance of this body of theoretical work is substantiated by a body of empirical evidence found in a number of recent papers: Boissay and Gropp (2007), Jacobsen (2015), Barrot (2016), and Costello (2020) among others.

model of these externalities is Diamond (1982) where, via search, consumption decisions affect output. In most of the literature, demand externalities result from nominal rigidities. There has been a recent interest in coupling nominal rigidities with financial constraints—for example, Eggertsson and Krugman (2012) and Guerrieri and Lorenzoni (2017).[4] The nature of demand externalities here is different. In particular, the type of expenditures by the private or public sector matters: spot orders may stimulate output, but chained orders depress it. The demand externalities arising from payment slowdowns are part of the classic narrative of payment slowdowns.

2 Empirical Motivation

Townsend's Turnpike: A prelude. To set the stage for our empirical analysis, we recall the *Townsend's Turnpike* (see Townsend, 1983). Townsend Turnpike is a representation of sequential transactions that occur in a variety of models where, unlike Walrasian settings where payments are instantaneous, the timing and direction of payments matter, (Ljungqvist and Sargent, 2018, ch. 28). Figure 1 presents a graph corresponding to a segment of Townsend's Turnpike. A subset of nodes in purple, $\{S_1, S_2, S_3\}$, represent agents endowed with funds who place orders for goods—i.e., agents with *spot funds*. To clarify terminology, we refer to spot funds as any medium of exchange available for immediate payment. Practical examples include cash accumulated from internal savings, business credit lines, and even trade credit. For the purpose of transferring goods, it is the same whether the customer pays with internal funds, borrowed funds from a bank, or IOUs to its suppliers. The critical feature of spot payments is that the delivery of goods is not contingent on other transactions. The remaining nodes, shown in green, $\{C_1, C_2, \ldots, C_6\}$, represent agents who obtain funds only after receiving payment; their ability to make payments is therefore *chained* to earlier transactions. Chained orders represent unfulfilled orders, contracts that promises to deliver goods only upon payment at future date.

The direction of each arrow indicates the flow of payments (i.e., who pays whom or, equivalently, who buys from whom). In this graphical example, the spot nodes pay their suppliers at time $\tau(0)$. Only then can the chained nodes that supply a spot node pay their own suppliers, at time $\tau(1)$. The sequence of payments propagates down-

[4] Recent papers have further introduced sequential transactions into environments with nominal rigidities—for instance, Woodford (2022) and Guerrieri et al. (2022). In those models, demand externalities occur when agents cut back on any form of expenditure.

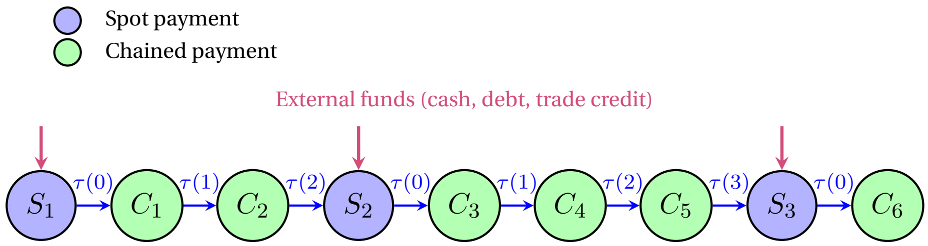


Figure 1: Sequence of Spot and Chained Payments

Notes: Purple nodes (S_1, S_2, S_3) hold external funds and initiate spot payments at $\tau^{(0)}$; green nodes are paid sequentially as funds propagate.

stream, forming a *payments chain* that ends when a chained node purchases from a spot node—the spot's payment is not contingent on receiving funds from others. In this example, the *chain length* n equals the number of downstream links reached by the initial funds.[5] The contractor example in the introduction fits this description, as do environments with geographic separation, random search, or production networks with sequential trade (e.g., Townsend, 1983; Freeman, 1996b; Kiyotaki and Wright, 1989; Kiyotaki and Moore, 1997a). By contrast, the payment by and good delivery to chained nodes are contingent on earlier transactions.

This section provides evidence on the existence and economic significance of payment chains, using detailed data on business-to-business transactions. First, we document the presence of chained payments—payments contingent on receiving earlier payments. Second, we show that firms engaging in chained payments are more likely to be financially constrained and to experience economic losses due to insufficient liquidity for upfront payments. Third, we construct payment chains from the data to characterize the structure of the network. Finally, we examine a financial disruption in which a subset of firms lost access to their funds. We show that these firms became more likely to engage in chained payments and trace the resulting economic losses among their trading partners along the supply chain.

Data. We leverage two exceptionally rich data sources. The first is Sinpe—Costa Rica's National Electronic Payments System—which processes and settles every interbank transfer.[6] In 2024 alone, Sinpe recorded over 62 million transfers, capturing, for each transaction, the precise timestamp, unique sender and receiver identifiers, the banks

[5]In Figure 1, S_1 and S_2 inject funds into chains of lengths 2 and 3, respectively.

[6]Sinpe is an acronym for Costa Rica's National Electronic Payments System (*Sistema Nacional de Pagos Electrónicos*), and is operated by the Central Bank of Costa Rica.

involved, and the exact transferred amount. The coverage spans the universe of interbank activity for 243,234 distinct firms, with an average transaction of USD 4,507.[7] The second dataset derives from Costa Rica's VAT (Value-Added Tax) reporting system, and encompasses the electronic invoices for every formal sale. For each invoice, the registry records the sale price, product code, quantity sold, and flags whether payment was settled on the spot or was deferred. Deferred payments are recorded because, until the seller receives full payment, classifying the transaction as deferred allows her to postpone the VAT remittance on the unpaid portion for up to 90 days. For deferred payments, the system also records the contractual repayment window. This feature provides direct, time-stamped insight into short-term financing arrangements and payment timing at a level of granularity rarely available.

Documented Regularities. We exploit the universe of interbank transfers among firms to test for systematic payment–chain behavior. Absent such chains, there should be no systematic relationship between the payment inflows and outflows of a firm. Thus, to empirically detect chained payments, we first examine if firm j is more likely to issue a payment within 24 hours of receiving one. Specifically, we estimate:

$$sent_{jt} = \beta \, received_{j,t-1} + \lambda_t + \lambda_{j \times d} + \varepsilon_{jt}, \tag{1}$$

where $sent_{jt} = 1$, if firm j made a payment on business day t, and $received_{j,t-1} = 1$ if firm j received a payment in the preceding 24 hours.[8] To prevent this relationship from being purely mechanical, we include both time fixed effects and firm×day-of-week fixed effects.[9]

Table 1 reports the results. Column (1) shows that, on average, firms are significantly more likely to send a payment within 24 hours of receiving a payment. Reassuringly, payments are only triggered when the amount sent is *less* than the amount received, as shown in Table A.1. Columns (2)–(4) reveal heterogeneity: the payment–chain effect

[7] For more details related to Costa Rica's firm network, see Méndez and Van Patten (2025); Argente et al. (2025). Alvarez et al. (2023) focuses on a particular branch of Sinpe— Sinpe Móvil—which pertains to peer-to-peer mobile money transfers.

[8] Payments received on weekends or holidays are assigned to the next business day.

[9] Time fixed effects capture calendar-driven patterns—e.g., firms that always pay or receive on month-end or the 15th. Firm × day-of-week fixed effects absorb idiosyncratic weekly rhythms—e.g., some firms systematically pay on Thursdays, others on Fridays—thereby isolating only the abnormal deviations linked to incoming payments. The latter fixed effect would also discipline variation, for instance, from a firm that tends to make payments every day, and instead would force the variation to come from out-of-the-ordinary responses correlated with receiving a payment.

is strongest among small firms (5 workers or less) and attenuated for firms with higher asset-to-sales ratios or greater loan access relative to sales.

Table 1: Triggered Payments Across Firms

Dependent variable: Probability that firm j makes a payment at time t

	(1)	(2)	(3)	(4)	(5)	(6)
$received_{j,t-1}$	0.030	0.010	0.039	0.038		0.030
	(0.0005)***	(0.0008)***	(0.0008)***	(0.0008)***		(0.0005)***
$Small_j \times received_{j,t-1}$		0.017				
		(0.001)***				
$\frac{Assets_j}{Sales_j} \times received_{j,t-1}$			-0.019			
			(0.001)***			
$\frac{Liabilities_j}{Sales_j} \times received_{j,t-1}$				-0.018		
				(0.001)***		
$received_{j,t-5}$					0.011	0.010
					(0.0004)***	(0.0004)***
Adjusted R^2	0.483	0.446	0.483	0.483	0.483	0.483
Observations	10,001,248	5,290,408	8,399,599	8,399,599	10,001,248	10,001,248
Clusters	222,306	74,566	174,896	174,896	222,306	222,306
$t, j \times$ d FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The independent variable in column (1) is an indicator that equals one if firm j received a payment from another firm in the last 24 hours. This variable is interacted with firm j's size, assets-to-sales, and liabilities-to-sales in columns (2), (3), and (4), respectively. Column (5)'s independent variable is an indicator that equals one if firm j received a payment within 5 days of time t. Column (6) includes both indicators described above. Robust standard errors, adjusted for clustering by firm, are in parentheses. We include time and firm×weekday fixed-effects. Data is daily and spans 2024.

Columns (1)–(4) focus on a one-day window to approximate chained-payment responses, as a narrow time frame strengthens the inference of a direct link between inflows and outflows. To trace the temporal profile of this effect, we implement a local projection (Jordà, 2005; Jordà and Taylor, 2025). Figure A.1 shows that, while most payments are triggered within one business day, there is a second peak around five business days (one week) after receipt of funds. Moreover, coefficients in columns (1), (5), and (6) are remarkably stable, suggesting that payments unfold in two orthogonal waves—an immediate response and a delayed one around business day five.[10]

This two-wave structure reflects both the speed of Costa Rica's modern payment infrastructure and residual frictions. In 2024, when Sinpe processes most transfers electronically, next-day responses are technologically feasible.[11] However, even with

[10] We also examine the intensive margin in Table A.2, which yields results consistent with the extensive-margin analysis: firms are more likely to make payments once they receive transfers.

[11] By contrast, in the 1990s, most payments settled by check, and checks used to require up to 15 days to clear—so that the fastest possible response was about two weeks (Cerdas and Melegatti, 2014).

a rapid payments network in place, a subset of firms continues to rely on paper checks, which take several business days to clear. This check-clearance lag produces the pronounced second spike at business day five. Consistent with this mechanism, firms exhibiting the five-day payment lag are roughly 60% more likely to pay by check than the rest. Moreover, when we exclude these check-using firms, the day-five peak of Figure A.1 disappears, confirming that check-clearing times underlie the longer delays.

Placebo Exercises Relations are not mechanical. We conduct placebo tests, where we change the timing so that the independent variable represents payments received h days *after* making a payment, with $h \in \{1, 2, 3\}$. As shown in Table A.3, results are consistently negative or insignificant, speaking against spurious temporal correlations.

Constructing Payments-Chains. Leveraging the comprehensive ACH and wire transfers data, we identify firms where the receipt of a payment triggers the issuance of subsequent payments. We classify firms into three distinct groups: (i) those that tend to make chained payments with a one-day delay; (ii) those that tend to make chained payments with a five-day delay; and (iii) those that tend to make spot payments.

We proceed in two stages. First, to isolate chained firms with a one-day delay (type (i)), we compute the monthly frequency with which receiving any payment leads to issuing a payment *within one business day* for each firm. As shown in Panel (a) of Figure A.2, this distribution has large masses of firms concentrated in a value of one (firms with perfect same-day response rates) and in a value of zero (firms with no same-day responses). Given this bimodal distribution, we use k-means clustering to split the sample into two groups: firms type (i) and other firms.

Second, among those firms that were *not* classified as type (i), and informed by the local projection in Figure A.1, we consider the distribution of monthly occurrences in which receiving a payment led a firm to make a payment five days later. This distribution, shown in Panel (b) of Figure A.2, shows that relatively few firms are chained with a five-day delay and allows us to further split this subsample into two groups: firms concentrated in zero, which correspond with type (iii)—spot, for whom payments sent are uncorrelated with payments received, and a few firms which are concentrated in one and are therefore type (ii). This two-stage clustering yields a firm-month classification into three types. Although firm types are highly persistent, our approach allows for type-switching from month to month.[12] Once firms are classified, we construct

[12]As an alternative, one can estimate each firm's fitted probability of triggered payments from the

payments-chain networks.

Stylized Facts on Payments-Chains Several stylized facts emerge. First, on a given month, 39% of firms tend to make spot payments, while the remaining 61% of firms engage mainly in chained payments. Among the chained firms, 80% make payments with a one-day delay and 20% make payments with a five-day delay.[13]

Thus, chained firms are prevalent in the economy, a prevalence likely driven by the generally short delays most firms incur. If payment delays raise the cost of participating in a chained network, it is natural to ask whether longer delays deter firms from becoming chained. To explore this, Figure 2 plots each sector's share of chained firms against its average payment delay. For instance, sectors in which firms tend to rely more on paper checks would have a longer average delay. The figure reveals a strong, negative relationship: sectors with longer delays host proportionally fewer chained participants.[14]

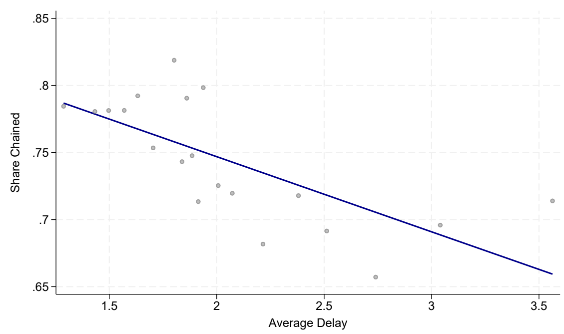


Figure 2: Chained Payments vs. Average Delays

Notes: The figure shows the average share of firms classified as chained by sector against the average payment delay in the sector. Sectors correspond with ISIC sections. Observations are weighted by sales. Data is monthly and spans 2024.

A second observation is that interactions across firm types are common. As shown in panel (a) of Figure 3, 26% of all links in the payments network are "mixed," involving both a spot and a chained firm.[15] Third, if we focus on transactions between same-type firms, network structures differ significantly. On the one hand, Panel (b) of Figure

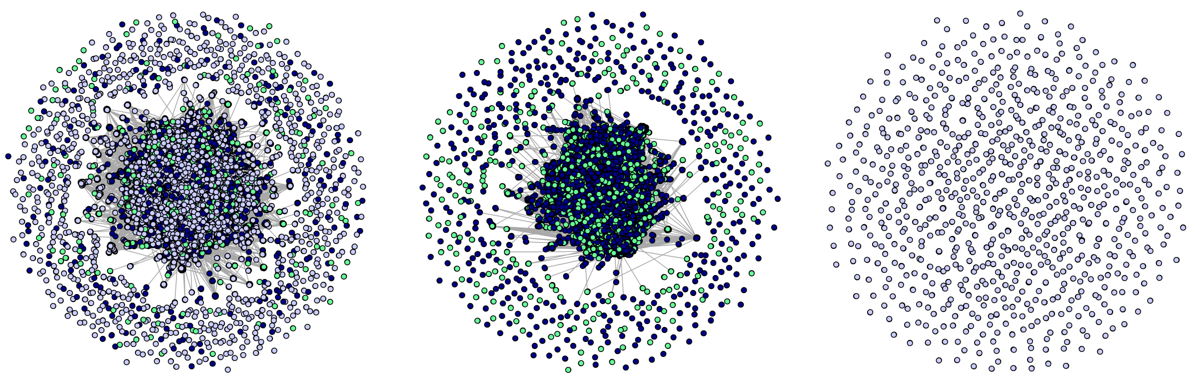
specification in column (6) of Table 1 and then apply k-means clustering with $k = 3$. The resulting spot vs. chained classification highly correlates with our baseline method, with a correlation of 0.89.

[13]The k-means clustering partitions firms at a fitted probability threshold of 0.42 when considering chained without delay vs. other firms, and of 0.15 when considering chained with delay vs. spot firms.

[14]Panel (b) of Figure 2 categorizes sectors by their one-digit ISIC sections. These broad groupings are chosen deliberately to have an input-output matrix with strongly dominant diagonals.

[15]This representation is based on December 2024. However, these shares are very stable: across all months in 2024, these shares are statistically equal and vary less than one percentage point from month to month, as shown in Table A.4.

3 shows the network for transactions between firms classified as chained, blue (green) dots representing firms chained without (with) delay; about 70% of all links. These links are particularly relevant as consecutive chained links ultimately form the payment chains. The network structure shown is a dense core with isolated nodes. The core cluster represents a critical mass of firms in tight transactional relationships and a circular dependency loop. Firms in this network's core are interdependent, which would relate to high systemic risk: if one core firm delays or defaults, it could ripple through the network. Finally, Panel (c) shows the transactions among firms classified as spot.[16] The network structure is more diffuse and decentralized.



(a) All (b) Chained-to-chained (c) Spot-to-spot

Figure 3: Network Structure by Firm Type

Notes: The figures show the network structure by link type. Panel (a) shows all links between firms. Panel (b) shows links between firms classified as chained. Panel (c) focuses on links between firms classified as spot. Data is collapsed for December 2024 to construct the network representation.

With firm classifications in hand, we construct payment chains as network paths that begin and end at firms making spot payments, with chained firms potentially occupying the intermediate nodes. For each pair of spot firms, we define the chain's length as the number of links (chained firms) separating them.[17] Figure A.3 displays the distribution of chain sizes with an average chain of 5 links.[18] The latter implies that, on average, the last firm in a chain suffers a delay of between 5-25 business days a year, corresponding with a delay and productivity loss of between 2% and 10% per

[16]These circles look less numerous than in Panel (a), as spot firms trade mostly with chained firms.

[17]To avoid double counting of overlapping fragments and to maintain computational feasibility at a monthly frequency, we measure chain length by the shortest path between two spot firms using Dijkstra's algorithm.

[18]This figure corresponds with the distribution for December 2024. Results are nearly identical in other months, exhibiting only minimal month-to-month variation.

year, with longer chains being associated with longer delays.

We further exploit the VAT registry of electronic invoices—linked at the firm level to Sinpe transfers—to observe prices and timing of payment—immediate or deferred, and if deferred, for how long—for firms classified as either chained or spot. Because sellers must issue an electronic invoice upon receipt of any partial payment, deferred transactions in the VAT system naturally capture chained-payment behavior. Indeed, firms classified as chained are 26% more likely to engage in deferred payments than spot-payment firms, with an average deferral of 34 days. Leveraging this rare level of detail, we find that, for the same product, invoice prices are 21% higher when payment is deferred, and that this premium is even larger when the purchasing firm is one that regularly makes chained payments. Appendix A.5 presents details of these results.

We also leverage daily data on deferred payments and itemized invoices to study how payment timing affects product delivery. Focusing on retail firms—where the items sold correspond exactly to the items purchased without any in-house transformation —allows us to match each input purchase directly to its downstream sale.[19] For these firms, we isolate new-product transactions with deferred payment and ask: does the retailer start selling the product as soon as the invoice is issued (potentially when receiving a partial payment), or only once the full payment arrives (recall the average deferral is 34 days)? In roughly 70% of cases, sales of the product start only *after* the upstream seller has received payment in full. The timing of such transactions is shown in Figure A.4—an event study centered on the date when *full* payment is due, where the plotted coefficients measure store-level sales (our proxy for delivery).

A Payments-Chain Crisis. A payments-chain crisis may arise when shocks to short-term liquidity propagate through chained transactions and impair downstream payments. We exploit a natural experiment to assess the empirical relevance of such crises.

In August 2024, Costa Rica's *Financiera Desyfin*—a financial institution catering mainly to enterprises and business clients—was abruptly and unexpectedly intervened by the financial regulator, who immediately froze all customer accounts due to insolvency concerns. The regulator then declared the institution unviable and placed it into a resolution process under the Organic Law of the Central Bank of Costa Rica, leaving those funds inaccessible for several months.

[19] For example, an automaker buys tires and parts but transforms them into cars, so input and output dates don't align neatly.

Using Sinpe data, we can identify the directly shocked firms and trace how disruptions ripple through their surrounding payment networks. To build intuition, we provide a simple example. Consider the pre-freeze network in Figure 1. Suppose firm S_2 held its funds at Desyfin. Before the freeze, this firm operated as a spot payer: it issued transfers to chained firms 3, 4, and 5 downstream, with no systematic link between the timing of its incoming and outgoing payments. After the Desyfin freeze, however, suppose S_2's outflows began to occur only after it received a payment, so that it "switched" its behavior to that of a chained firm. This shift would delay the timing of payments to firms 3, 4, and 5 downstream, lengthening the time they had to wait to receive funds.

With this example in mind, let $Desyfin_i = 1$ if firm i maintained an account at Financiera Desyfin during 2024, and let $Shock_t = 1$ for $t \geq$ August 2024, when the freeze occurred. We also define $Switch2chained_{it} = 1$ if firm i changes from being classified as spot to being classified as chained in month t.[20]

We first explore if firms which had an account at Desyfin become more likely to "switch" to being classified as chained after the freeze. Column (1) of Table 2 reports that this occurs; the probability of switching from spot to chained increased by 0.016 pp—equivalent to a 9% rise relative to the baseline.[21] The estimation has province×industry and time fixed effects, and is therefore not driven by regional or sectoral composition or aggregate shocks that year.

Next, we examine how a "switch" to chained status impacts a directly shocked firm (D, in our example above). Namely, we consider the indicator $Switch2chained_{it}$ as an independent variable and study the behavior of a firm i that had an account at Desyfin *and* switched to being classified as chained. We find that firm i made 18% fewer purchases after the shock, which would align with a hindered ability to make payments, as shown in column (2) of Table 2. The same column shows that the term $Desyfin_i \times Shock_t$ has, on its own, a positive coefficient, consistent with positive pre-shock selection into Desyfin, which catered to corporate clients. In line with the decline in purchases, column (3) shows that sales fell, and column (4) confirms that the marginal effect on value added is a 20% decrease.

We finally ask how chained suppliers fare when selling to a directly shocked firm. Define $Selling2Desyfin_{jt}$ as an indicator equal to one if chained firm j was *selling to* a

[20] Recall that the classification of firms by type is month-specific and allows for type-switching across time. While the classification is flexible, firm types tend to be highly persistent.

[21] This result is also reassuring in that it validates our classification measure as capturing meaningful payment behavior.

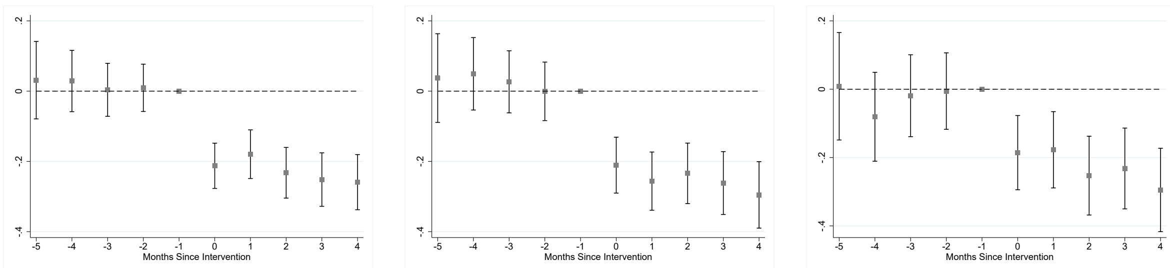
Table 2: A Payments-Chains Crisis: Frozen Bank Accounts

A. Directly Shocked Firm i				
	$Switch2chained_{it}$	$\ln\ Purchases_{it}$	$\ln\ Sales_{it}$	$\ln\ Value\ Added_{it}$
	(1)	(2)	(3)	(4)
$Switch2chained_{it} \times Desyfin_i \times Shock_t$		-0.176	-0.233	-0.197
		(0.064)***	(0.086)***	(0.109)***
$Switch2chained_{it} \times Shock_t$		-0.004	0.019	-0.017
		(0.577)	(0.010)**	(0.012)
$Desyfin_i \times Shock_t$	0.016	0.159	0.195	0.204
	(0.007)***	(0.048)***	(0.037)***	(0.048)***
%Δ w.r.t. mean [mean dep. var]	9% [0.18]			
Observations	1,200,404	1,192,436	939,984	636,485
Province×industry, t FE	Yes	Yes	Yes	Yes

B. Firm j Selling to a Directly Shocked Firm i			
	$\ln\ Purchases_{jt}$	$\ln\ Sales_{jt}$	$\ln\ Value\ Added_{jt}$
	(5)	(6)	(7)
$Selling2Desyfin_{jt} \times Switch2chained_{it}$	-0.176	-0.189	-0.130
$\times Desyfin_i \times Shock_t$	(0.029)***	(0.033)***	(0.038)***
$Switch2chained_{it} \times Shock_t$	0.038	0.017	0.000
	(0.016)**	(0.019)	(0.023)
$Selling2Desyfin_{jt} \times Switch2chained_{it}$	0.651	0.669	0.661
	(0.040)***	(0.044)***	(0.047)***
$Desyfin_i \times Shock_t$	0.035	0.027	0.000
	(0.016)**	(0.019)	(0.023)
Observations	704,087	555,098	382,319
Province×industry and t FE	Yes	Yes	Yes

Notes: The dependent variable in column (1) is an indicator equal to 1 if firm i switched from being classified as a spot firm to a chained firm in period t. Columns (2), (3), and (4) investigate the impact of the freeze on firms who held accounts at Desyfin, before and after the freeze, differentiating between impacts if the firm switched to being classified as chained or not. Columns (5), (6), and (7) instead focus on firms that were selling to a firm directly connected to Desyfin, before and after the shock. Robust standard errors, clustered by firm, are in parentheses. We include province×industry (ISIC-4) and time fixed-effects. Data is monthly and spans 2024.

directly shocked "Desyfin" firm. Panel (B) of Table 2 shows an 18% reduction in purchases after the freeze, along with a fall in sales and value added of a similar magnitude (columns 5-7), *but only when selling to a firm that switched to chained payments and held a Desyfin account.* These results underscore how payment chains can undermine economic resilience. Just as a project developer's lack of funding can trigger cascading delays for the general contractor, subcontractors, and so on, a shock—such as a bank account freeze—that disrupts an initial spot payment can cause delays and hinder production for firms down the payments chain. This effect on "one-link away" firms is persistent. While none of these firms was directly hit by a shock, Figure 4 shows that the effect remains at least four months after the shock occurred—the figure again focuses on firms selling to a Desyfin firm *which switched* to being classified as chained; i.e., we plot the interaction $Selling2Desyfin_{jt} \times Switch2chained_{it} \times Desyfin_i \times Shock_t$. The effect is also persistent along a chain. As documented in Figure A.6, the effect propagates at least 3 links upstream from a directly hit Desyfin firm that switched to chained.[22]



(a) $\ln Purchases_{jt}$ (b) $\ln Sales_{jt}$ (c) $\ln Value\ Added_{jt}$

Figure 4: Persistent Effect Along the Chain

Notes: The figure shows an event study where time zero corresponds with the month when Desyfin accounts were frozen. The dependent variables in panels (a), (b), and (c) are the (log) of purchases, of sales, and of value added, respectively, of firm j in month t. The plotted coefficients correspond with the triple interaction of $Selling2Desyfin_{jt} \times Switch2chained_{it} \times Desyfin_i \times Shock_t$, and therefore captures the impact of being one link upstream from a Desyfin firm *and* switching to being classified as chained *after* the freeze. Robust standard errors are clustered by firm. We include province×industry (ISIC-4) and time fixed-effects. Data is monthly and spans 2024.

3 Payment-Chains and Productivity

This section presents a framework to map a set of orders representing expenditures paid on the spot or not, into a payment-chain network. When transaction delays entail an economic loss, the payments-chain network translates into TFP losses. We then show how to aggregate a payment-chain network into a portable budget constraint with externalities.

Bilateral Relations. Consider an economy where production is organized through bilateral trade. As in our introduction, we can think of these inputs as construction materials or services and, more generally, as specially engineered products. For simplicity, we assume that all input goods are identical and, thus, the numeraire. The price paid and, thus, the transaction amount are identical across orders.

[22] The sample size becomes smaller as we increase the number of links of separation, because fewer observations feature chains that long (see Figure A.3). Thus, results beyond 3 links are too noisy to be conclusive.

There are two types of orders: *spot* and *chained* orders. Spot orders are paid immediately. Chained orders are contingent on previous payments; the customer pays for his order, contingent on receiving payment from another transaction in which he acts as a supplier. There are N^s and N^x spot and chained orders. The sum of orders equals the amount N of input goods, $N = N^s + N^x$, a notion of market-clearing that will emerge in equilibrium in the next section. We provide an interpretation for spot and chained transactions once we introduce financial decisions in that section.

Each input unit is assigned an identifier, $i \in \mathcal{N} = \{1, 2, \ldots, N\}$. Likewise, each order is assigned a unique identifier, $i \in \mathcal{N}$. We partition $\mathcal{N}$ into two sets, $\mathcal{N}^s$ and $\mathcal{N}^x$, denoting the set of identifiers of spot and chained orders, respectively.[23]

We define two relations that together determine a payment-chain network. First, $\mathcal{P} : \mathcal{N} \to \mathcal{N}$ is a one-to-one assignment from an order to an input satisfying $i \neq \mathcal{P}(i)$. $\mathcal{P}(i)$ indicates the input unit that fulfills order i. Second, a chained income-expenditure relation associates a chained order with an input unit whose sale is the source of funds of an order of another input. This relation is the identity function defined on $\mathcal{N}^x$.[24] That is, order i uses the payments to input unit i if $i \in \mathcal{N}^x$, as a source of funds to execute order i. The idea is that although order i is not externally funded, the customer who places order i owns input unit i. Hence, order i can be funded after order $j = \mathcal{P}^{-1}(i)$ executes a payment.

Initially, we do not impose structure on the assignment $\mathcal{P}$. Different specifications of $\mathcal{P}$, e.g., random assignments or assignments requiring some form of economic organization, lead to different network structures and, hence, different predictions. To understand how these relations induce a payment-chain network, consider input unit j assigned to order i, i.e., $\mathcal{P}(i) = j$. The client placing order i must pay for input j. This creates a payment link from i to j. In turn, if order j is spot, $j \in \mathcal{N}^s$, the funds used to pay for order i are not used to make further payments because j is funded from an external source of funds. However, if order j is chained, $j \in \mathcal{N}^x$, the funds are used again to pay for input $k = \mathcal{P}(j)$. In other words, when $j \in \mathcal{N}^x$, there is a flow of payments from i to j, and from j to k. If order k is also chained, the same funds continue to flow and pay

[23] $\mathcal{N}^s$ and $\mathcal{N}^x$ are a partition of $\mathcal{N}$: $\mathcal{N}^s \cap \mathcal{N}^x = \emptyset$ and $\mathcal{N}^s \cup \mathcal{N}^x = \mathcal{N}$ with $N^s = |\mathcal{N}^s|$, $N^x = |\mathcal{N}^x|$.

[24] The assumption that the income-expenditure relation is the identity is without loss of generality. Formally, we can define the chained income-expenditure relation $\mathcal{X}$ as the identity function on $\mathcal{N}^x$, that is $\mathcal{X} : \mathcal{N}^x \to \mathcal{N}^x$ such that $\mathcal{X}(i) = i$. The idea is that $i \in \mathcal{N}^x$ obtains funds from input unit $\mathcal{X}(i) = i$. Indeed, the identity function $\mathcal{X}(i)$ can be replaced by any injective function $\mathcal{X} : \mathcal{N}^x \to \mathcal{N}$ so that input units and associated chained orders do not have the same identifiers. Changing identifiers does not affect the results.

input $\mathcal{P}(k)$, and so on. The payments chain continues until a final order in sequence of payments is placed on some input unit i that is spot. Since every order is paired with an input unit, the economy features an entire network of transactions forming a collection of payment chains. The network captures the economy-wide transactions described in the empirical analysis, albeit simplifying by assuming payments are of equal size and do not have to confront size incongruence.

The payment network would be inconsequential if every individual payment were executed immediately because the transfer of funds through the entire network would be instantaneous in that case. In the extreme, a single dollar could finance an entire economy's transactions if passed on and settled immediately. Yet even if transfers are not instantaneous, the network would be inconsequential if no resources were lost due to payment speed. As documented in our empirical analysis, we observe transaction delays, dependence on income flows to execute expenditures, and non-trivial economic losses for those who don't pay for orders immediately. When there are economic consequences from delays, the payments-chain network determines who and by how much bears those losses.[25]

To formalize the notion of delays and losses, we define an order's **position**, n, in a payment chain as the number of payments that must be executed before an order is funded. Naturally, spot orders occupy position $n = 0$, chained orders receiving payment from spot orders occupy position $n = 1$, chained orders receiving payment from orders in position $n = 1$ occupy position $n = 2$, and so on. To capture the timing of payments, let $\tau(n)$ denote the time at which payment is executed for position n, with $0 = \tau(0) < \tau(1) < \tau(2) < \cdots < 1$. Transaction delays can arise due to limited technology to process transactions, e.g., the reliance on physical currency transfers or the clearing time of checks and electronic payments. They may also serve as a protocol for withholding funds for some time to prevent fraud. To capture the economic losses from delays, let y_n denote the amount of final goods received by an order at position n, where $y(0) = 1 \geq y(1) \geq y(2) \geq \cdots > 0$. This degradation can arise from depreciation or reduced production times. In what follows, we first analyze the distribution of chain lengths and aggregate implications, without reference to $\mathcal{P}$, τ, or y. We then consider two special cases of $\mathcal{P}$ and provide two economic motivations for choices of τ and y.

[25] Recall that in the empirical analysis, the last firm in the average chain was found to suffer a delay of between 5-25 business days a year, equivalent to a delay and productivity loss of between 2% and 10% per year, with longer chains being associated with longer delays.

Payments-Chain Network. Given the assignment $\mathcal{P}$ and the set of spot and chained orders, we construct a payment-chain network. A payment-chain network is the collection of payment chains, encompassing the universe of transactions. We define an individual payment chain as a sequence of payments starting from each spot order with external funding, followed by a sequence of chained orders in which those initial funds are transferred from one transaction to another, ending with a chained order that buys inputs that are not the funds for another chained order (the input with same identifier of a spot order). The number of chained orders defines the chain length—the length is zero if a spot order follows the initial spot order. Formally:

Definition 1. *A **payments-chain network** $\mathcal{K}$ is an acyclical directed network with nodes $\mathcal{N} = \{1, 2, ..., N\}$ and links $\mathcal{V} = \{(i, j)|\mathcal{P}(i) = j, j \in \mathcal{N}^x\}$, $\mathcal{K} = (\mathcal{N}, \mathcal{V})$. A **payments chain** of length n is a finite sequence of nodes $\{i_k\}_{k=0}^n$ such that the sequence starts with some $i_0 \in \mathcal{N}^s$ and $\forall k \in \{1, \ldots, n\}$, $\mathcal{P}(i_{k-1}) = i_k \in \mathcal{N}^x$. By convention, if $i \in \mathcal{N}^s, \mathcal{P}(i) = j \in \mathcal{N}^s$, $\{i, j\}$ form a chain of length zero.*

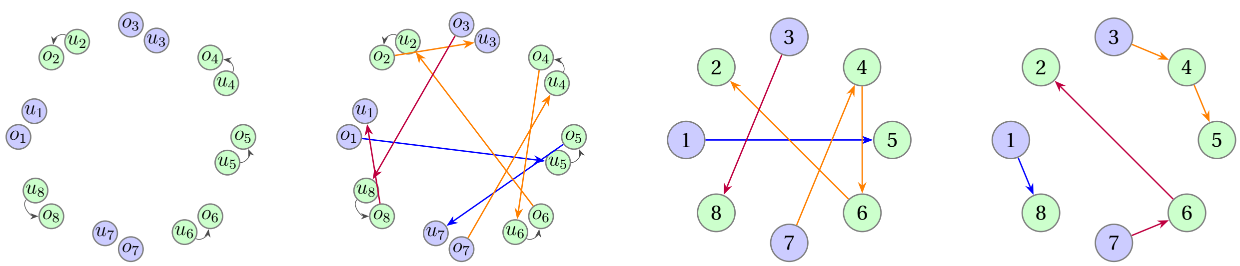
In the definition, nodes in the payment-chain network represent both orders and input units. The directed links represent the direction of flows of funds. As transactions are bilateral, at most one link stems from each node. A link from i to j indicates that i orders from input unit j and that j is a chained order, $\mathcal{P}(i) = j \in \mathcal{N}^x$. The source of funds for order j is the payment from order i. In turn, if a node does not receive a link, it is a spot order. If no link is directed toward order $j = \mathcal{P}(i)$, then order j is also spot—in which case we say that $\{i, j\}$ form a zero-length chain. In this construction, each order is guaranteed a source of funds, either internally via the network or externally. The assumption that the network is acyclic rules-out cycles where chained orders form closed loops and, thus, would be unfunded. Laying out the network in a straight line running to infinity, $N \to \infty$, would correspond to a Townsend turnpike.

Examples. Let us provide an example. Set $\mathcal{N} = \{1, 2, \ldots, 8\}$ and let the subset of spot orders be $\mathcal{N}^s = \{1, 3, 7\}$. Also, define the assignment relation as follows: let $\{i_n\}_{n \in \mathcal{N}} = \{1, 5, 7, 4, 6, 2, 3, 8\}$ such that $i_{n+1} = \mathcal{P}(i_n)$ and $i_1 = \mathcal{P}(i_N)$. Thus, the links in this payments-chain network are $\mathcal{V} = \{(1, 5), (7, 4), (4, 6), (6, 2), (3, 8)\}$.

Several graphs are associated with this example. Panel (a) of Figure 5 depicts the chained income-expenditure relation. In that panel, we split each node into counterparts: the input units $\{u_n\}$ and orders $\{o_n\}$ for nodes $n \in \{1, 2, ..., 8\}$. The links represent the flow of funds from input units to their corresponding chained orders, defined by the chained income-expenditure relation.

Panel (b) adds the links to the flow of payments for inputs, corresponding to $\mathcal{P}$—links in the panel add the payments to input units. Adding the links of the chained income-expenditure and assignment relations allows us to trace funds. Notice that the links from orders to input units with the same color share a common source of funds.

Panel (c) depicts the resulting payments-chain network. In this example, there are three chains with lengths $(1, 3, 1)$.[26] Note that this network generates unequal losses: order 2 at position $n = 3$ receives only y_3, while orders at positions $n = 1$ receive $y_1 > y_3$. The network could be reorganized as in Panel (c) to reduce these losses.[27] This observation motivates our efficiency analysis of assignments $\mathcal{P}$.



(a) Income relation (b) Network in detail (c) Random Network (d) Efficient Network

Figure 5: Components of the Payments-Chain Network

Chain-Length Distributions and Aggregate Output. We now derive moments of the payment-chain network as $N \to \infty$, while keeping the ratio $\mu \equiv N^x/N$ constant. Aggregate output per order depends on the distribution of positions that orders occupy. This position distribution is induced by the distribution of chain lengths, which in turn is determined by the assignment $\mathcal{P}$. Let $g(n; \mu, \mathcal{P})$ denote the probability mass function of chain length under assignment $\mathcal{P}$, and let $G(n; \mu, \mathcal{P}) = \sum_{m=0}^{n} g(m; \mu, \mathcal{P})$ denote its cumulative distribution function. Since position n is found only in chains of length n or greater, the fraction of orders at position n is proportional to $1 - G(n-1; \mu, \mathcal{P})$ leading to a distribution of positions:

$$H(n; \mu, \mathcal{P}) = \frac{1 - G(n-1; \mu, \mathcal{P})}{\sum_{k=0}^{\infty}[1 - G(k-1; \mu, \mathcal{P})]} \tag{2}$$

[26] The first chain from 1 to 5 (length 1), the second chain from 7 through 4, 6, to 2 (length 3), and the third chain from 3 to 8 (length 1).

[27] For instance, reassigning the links to create chains of lengths $(1, 2, 2)$ would improve output for order 2 (from y_3 to y_2) while only slightly reducing output for one other order.

where $G(-1) \equiv 0$. Hence, aggregate output per order is given by:

$$\mathcal{Y}(\mu, \mathcal{P}) = \sum_{n=0}^{\infty} H(n; \mu, \mathcal{P}) \cdot y_n = (1 - \mu) + \mu \cdot \mathcal{A}(\mu, \mathcal{P})$$

where $\mathcal{A}(\mu, \mathcal{P}) \equiv \mathbb{E}[y_n|\text{chained order}]$ is the expected output among chained orders. The key insight is that different assignments $\mathcal{P}$ generate different distributions $g(n; \mu, \mathcal{P})$ and thus different aggregate outcomes, given μ.

For now, examine two types of assignments that span a spectrum of possible network structures, leaving other variations to the extensions below. On one end, we have a **random assignment** where any input unit can be matched to any order with equal probability—serving as a benchmark representing decentralized markets without coordination. There are natural ways to construct a random assignment $\mathcal{P}$ that induces our payment-chain network. One approach treats $\mathcal{P}$ as generated by a random ordering of $\mathcal{N}$.[28] An alternative is to treat $\mathcal{P}$ directly as a permutation without fixed points—a derangement.[29] This alternative potentially induces cycles containing only chained orders—a phenomenon of possible interest but outside of our definition—but these problematic cycles will contain a measure-zero set of orders as $N \to \infty$, the limiting case of interest in this paper.[30] On the other end, the **efficient assignment** that maximizes $\mathcal{A}(\mu, \mathcal{P})$ given μ.

For the output function, we consider the exponential form $y_n = \delta^n$ where $\delta \in (0, 1)$. Appendix B.1presents possible interpretations of y_n. First, it can represent time decay: if $\tau(n) = n$ represents uniform time, δ^n captures discounting or depreciation. Alternatively, the function arises when intermediate products require time to be transformed into final goods, and payments are made after an inspection time proportional to the final output. While waiting for delivery, other inputs such as labor or energy remain idle.[31] In that case, $\tau(n) = 1 - \delta^n$. This latter interpretation is convenient as it ensures

[28]Draw a random ordering $(\pi(1), \ldots, \pi(N))$ of the elements. Define the assignment as $\mathcal{P}(\pi(i)) = \pi(i + 1)$. This creates a single production chain passing through all N nodes, and the payment-chain network is constructed by severing the links at spot orders. No cycles of chained orders can arise.

[29]For each node i, assign a supplier $\sigma(i)$ such that σ is a bijection with $\sigma(i) \neq i$ for all i.

[30]Shepp and Lloyd (1966) show that, for N large, the number of cycles of length k in a random permutation is asymptotically Poisson with parameter $1/k$. Thus, the expected number of nodes in k-cycles is $k \times \frac{1}{k} = 1$. For a cycle to be problematic, all elements must be chained orders, which occurs with probability μ^k. Hence, the expected number of nodes in problematic k-cycles is μ^k, and the total expected number of nodes in problematic cycles is $\sum_{k=2}^{\infty} \mu^k = \frac{\mu^2}{1-\mu}$, a finite constant. The fraction of such nodes therefore vanishes as $N \to \infty$.

[31]The mechanism works as follows: an order at position n receives inputs at time τ_n and produces

all transactions are completed within a unit time interval and is motivated as an escrow mechanism that prevents two-sided fraud in environments with customized goods.

Under either interpretation, the idleness of resources translates into TFP losses that depend on the position in a chain. For the two assignments and output function in consideration the distribution of chain lengths and TFP aggregate as follows:

Proposition 1 (Random versus Efficient Networks)**.** *Given a fraction μ of chained orders and output function $y_n = \delta^n$, as $N \to \infty$:*

Assignment	**Auxiliary Objects**	**Distribution** $g(n;\mu,\mathcal{P})$	**Chained Output** $\mathcal{A}(\mu;\delta,\mathcal{P})$
Random	–	$(1-\mu)\mu^n$	$\delta\frac{1-\mu}{1-\delta\mu}$
Efficient	$\underline{n} = \lfloor \mu/(1-\mu) \rfloor$ $\overline{n} = \lceil \mu/(1-\mu) \rceil$ $p = \overline{n} - \mu/(1-\mu)$	$\begin{cases} p \text{ if } n = \underline{n} \\ 1-p \text{ if } n = \overline{n} \\ 0 \text{ otherwise} \end{cases}$	$\delta\frac{1-[p\delta^{\underline{n}}+(1-p)\delta^{\overline{n}}]}{1-\delta}\frac{1-\mu}{\mu}$

$\mathcal{A}$ is concave and decreasing in μ and convex and increasing in δ with limits:

$$\lim_{\mu\to 0}\mathcal{A}(\mu;\delta) = \delta, \quad \lim_{\mu\to 1}\mathcal{A}(\mu;\delta) = 0, \quad \lim_{\delta\to 0}\mathcal{A}(\mu;\delta) = 0, \quad \lim_{\delta\to 1}\mathcal{A}(\mu;\delta) = 1.$$

The geometric distribution emerges naturally in random networks from the matching process. Consider building a payment chain starting from a spot order. This order is matched to an input unit. The chain continues if and only if this unit corresponds to a chained order, which occurs with probability μ. If it is chained, it links to another unit that is also chained with probability μ, and so on. The chain terminates when a spot order is reached, occurring with probability $(1-\mu)$. Thus, the probability of a chain having length n is $g(n;\mu) = (1-\mu)\mu^n$ as the distribution of positions, $H(n;\mu)$. The formula for output follows immediately from $y_n = \delta^n$.[32]

The efficient assignment minimizes output losses by equalizing chain lengths as much as the integer nature of the chains permits. To see why this is optimal, consider any assignment with chains of varying lengths. If we have chains of lengths n_1 and n_2 such that $n_1 + 1 < n_2$, we could move an order from position n_2 in the longer chain to

$y_n = 1-\tau_n$ using production technology $\int_{\tau_n}^1 \min\{x_t, h_t\}dt$ with $h_t \leq 1$ where h_t are labor resources. To prevent seller-side fraud, the buyer must inspect fraction $(1-\delta)$ of output before releasing escrowed payment, taking $(1-\delta)(1-\tau_n)$ time. This yields $\tau_{n+1} = \tau_n + (1-\delta)(1-\tau_n) = 1-\delta^{n+1}$, hence $y_n = \delta^n$.

[32]Note $\mathcal{Y} = \sum_{n=0}^{\infty} H(n;\mu)\cdot y_n = \sum_{n=0}^{\infty}(1-\mu)\mu^n \cdot \delta^n = \frac{1-\mu}{1-\mu\delta}$.

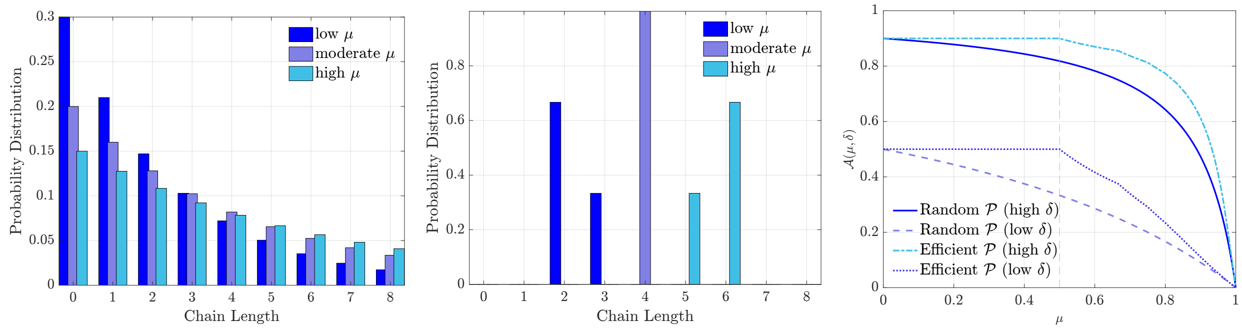
extend the shorter chain to length n_1+1. This reallocation strictly increases total output since the relocated order gains $y_{n_1+1} - y_{n_2} = \delta^{n_1+1} - \delta^{n_2} > 0$. Such improvements are possible whenever chains differ by more than one in length. Therefore, the efficient assignment must concentrate all chains on at most two consecutive integers: chains of length $\underline{n}$ and $\overline{n} = \underline{n} + 1$. The weight p on chains of length $\underline{n}$ ensures that the average ratio of spot to chained orders across chains yields the ratio across all orders $\mu/(1-\mu)$.

Under both assignments, $\mathcal{A}$ is concave and decreasing in μ with well-behaved limits.[33] $\mathcal{A}$ plays a critical role once orders result from agent decisions; concavity is an important property.

Figure 6 illustrates the key differences between random and efficient assignments. Panel (a) shows the geometric distribution of chain lengths following random assignment for three values of μ. Greater μ induces longer chains. Panel (b) shows the efficient assignment counterpart exhibiting distributions concentrated on only two adjacent lengths. Panel (c) compares $\mathcal{A}(\mu)$ across both assignments for $\delta \in \{0.5, 0.9\}$. Output per chained order decreases in μ, but the efficient assignment substantially outperforms the random assignment, with the gap widening as μ increases. Notice that under the efficient assignment for $\mu \leq 1/2$, $A(\mu) = \delta$ because it is possible to assign each chained order to a spot order.

The main insight from this section is that economic losses can result from payment delays and the possible disorder of payments. It is worth distinguishing the sources of these losses from other environments. Unlike search models, there are no congestion externalities—the production assignment is one-to-one. Unlike sticky-price models, prices have do not have allocative implications as long as goods are transferred. Instead, economic losses arise exclusively from delays, which is clear because the economy is at full capacity when all transactions are spot or $\delta \to 1$. This is true under both assignments of payments, but importantly, losses are magnified where payments are disorganized by the random assignment of orders. The efficient assignment minimizes these losses by concentrating all chains at similar lengths.

[33]As $\mu \to 1$, chains become arbitrarily long and the average final goods received per chained order tends to zero. Conversely, as $\mu \to 0$, a chained order is guaranteed to end in position 1, obtaining $\mathcal{A} = \delta$ in expected goods. Similarly, as $\delta \to 0$, chained orders receive nothing, whereas any losses vanish as $\delta \to 1$.



(a) g(n) (random assignment) (b) g(n) (efficient assignment) (c) Output $\mathcal{A}(\mu)$

Figure 6: Chain-Length Distributions and Average Output per Chained Order

Notes: Panels (a) and (b) show distributions $g(n;\mu)$ for $\mu \in \{0.7, 0.8, 0.85\}$ under random and efficient assignments. Panel (c) shows $\mathcal{A}(\mu)$ for both assignments with $\delta \in \{0.5, 0.9\}$.

3.1 Interpretations and Extensions

Economic Mechanisms for Network Formation. Random and efficient assignments analyzed above can arise from different market structures and environments. Random assignments arise naturally in centralized markets, where, as we show below, a Walrasian mechanism equates supply and demand for inputs but does not organize orders. In such a market, agents placing chained orders trade with uncertainty about the timing of inflows and the delivery of goods. This randomness may also result from the randomness of who wants to buy from whom. In turn, an efficient assignment emerges if there exists a market entity or mechanism that can organize orders and payments, thereby minimizing payment times. Below, we discuss how an entity that centralizes payments aggregates as a production function. The efficient assignment requires substantial information and coordination, which is not necessarily the most realistic assumption.

Alternative Assignments and Empirical Patterns. The evidence from Section 2 suggests that real payment networks lie somewhere between the polar extremes of random and efficient networks: The observed distribution is bell-shaped and concentrated around chains of length 4-5, unlike the geometric distribution that emerges from randomness. However, the empirical distribution shows a substantial share of zero-length, which indicates inefficiency.[34] These patterns motivate us to study alternative assignments in Appendix C.

[34] Recall that the efficient size distribution is concentrated in two contiguous integers.

A first extension explores **correlated assignments**, where links between orders of different type do not have i.i.d. probabilities of being linked. Instead, the likelihood of a link between chained orders is $\mu + \rho$ leading to a geometric distribution of chain lengths with parameter $\mu + \rho$, and efficiency $\mathcal{A}(\mu; \delta, \rho) = \delta \frac{1-(\mu+\rho)}{1-\delta(\mu+\rho)}$.[35]

We further consider a **pooled assignment**, a hybrid between random and efficient allocations where chained orders can be reorganized efficiently between chains but spot-to-spot links cannot. The formula for $\mathcal{A}$ matches that of the efficient assignment, but replacing the ratio of chained to spot orders, $\mu/(1-\mu)$, with the ratio of chained to available spot orders $1/(1-\mu)$. This pooled assignment captures income pooling the same agent places many chained orders. The relevance of this assignment becomes clear when we aggregate budgets.

Another possibility is costly **reassignment**, another hybrid between random and efficient chains. This assignment arises when orders are initially assigned randomly, but chain orders can be reallocated at a cost Ξ, thought of as negotiation costs. A threshold position η emerges endogenously from the trade-off between reassignment costs and losses from a position. The resulting distribution exhibits efficient clustering at some position m and then a geometric distribution up to η.

Variations to the environment. We consider two further extensions that do not modify the assignment, but the output function y_n. In one these extensions, we allow for **default and cascading**, where orders at position $n > \eta$ default, keeping funds rather than executing payments. This captures situations where the losses from delays exceed the opportunity cost of not obtaining goods. Once a transaction fails, subsequent transactions also fail due to a lack of funds. In this case, $y_n = 0$ for positions above the default threshold.[36]. This extension is interesting because, unlike the reassignment extension, which requires information, defaulting is a unilateral decision that only requires knowing the passage of time and can also increase efficiency.

Finally, we consider **delay heterogeneity**. In this extension, within any chain of length n, we allow the first k transactions to occur without delay. If k is drawn from a

[35] We show this by exploiting that the payment-chain network has a Markovian representation. When chained orders are less likely to connect with chained orders, $\rho < 0$, the economy achieves greater efficiency than under baseline i.i.d. random assignment, though falling short of full efficiency.

[36] If the chain-length distribution remains geometric, the resulting TFP is the baseline TFP scaled by $(1 - \delta^{\eta} p^{\eta})$

binomial with parameter q and n draws, the resulting TFP formula decomposes into:

$$\mathcal{A}_q = (1-q) + \delta \frac{q(1-\mu)}{1-\mu(1-q+\delta q)}. \tag{3}$$

When $q = 0$ (all instant), $\mathcal{A}_q = 1$; when $q = 1$ (all delayed), we recover the baseline formula.[37] This extension captures payment environments where instant payments coexist with traditional systems, e.g., checks. The extension is also meant to capture the possibility of distinguishing between trade credit and spot transactions along a supply chain (supply-chain finance).[38]

Finally, we generalize the formula for y_n to allow for **alternative uses** of resources while products are not delivered. In the microfounded setting for y_n, we assume that labor remains idle until inputs are delivered. Suppose instead that labor has a less efficient alternative use, producing $(1-\alpha)$ units of final goods per unit of time *before* inputs arrive and one unit of final goods per unit of time *after* inputs arrive at τ_n. Then, y_n is a weighted average of the outputs per unit of time, weighted by the duration when inputs are unavailable versus available:

$$y_n = (1-\alpha)(1-\delta^n) + \delta^n = 1 - \alpha + \alpha\delta^n.$$

Using the same aggregation properties and the linearity of expectations, aggregate output becomes:

$$\mathcal{Y} = (1-\mu) \cdot 1 + \mu \cdot \mathbb{E}[1-\alpha+\alpha\delta^n | n \geq 1] = (1-\alpha\mu) + \mu\alpha \cdot \mathcal{A}(\mu, \delta).$$

The parameter $\alpha \in [0,1]$ captures the productivity loss from not having the specialized input: $\alpha = 1$ recovers our baseline (complete idleness), while $\alpha = 0$ implies no productivity loss (perfect substitutes).

[37]The first term $(1-q)$ is the contribution from instant payments, while the second term captures the reduced productivity from delayed transactions.

[38]We treat spot orders as including trade credit. However, there is a nuanced distinction when the supplier extending trade credit is itself financially constrained. In such cases, trade credit ceases to be equivalent to a spot transaction and instead becomes a hybrid. If the downstream firm lacks liquidity, it must wait for the upstream firm to sell before it can repay its own obligations. Trade credit thus accelerates production at one node of the chain but delays another. The heterogeneous chain is meant to capture this possibility. Bocola (2022) study situations where buyers and sellers are constrained but does not feature the delays we emphasize on this extension.

3.2 From Payments-Chain Networks to Portable Budget Constraints

We now demonstrate how the transactions and delays in the payment chain aggregate into reduced-form budget constraints that can be embedded into standard models.

Timing. We want to incorporate payment chains into dynamic models with discrete timing $t \in \{0, 1, 2, \ldots\}$. To avoid tracking the distribution of order statuses across periods, we focus on transactions that occur between periods. For that purpose, we specialize in the payment-chain network above, with transaction times $\tau(n) = 1 - \delta^n \in [0, 1]$, for $n \in \{0, 1, 2, \ldots\}$. This allows us to study models where payment-chain networks are created at the start of every discrete time period, with transactions occurring within each interval in between periods $\tau(n) \in [0, 1]$ where payments clear by the end of the period. As before, inputs are the numeraire.

We aggregate the budget constraint of a single agent, e.g., a household or a firm, that has y inputs to sell at the start period t. This endowment is composed of N units of goods, each of size y/N. The agent also begins the period with financial wealth W and has placed spot orders and chained orders, valued at E^s and E^x, for N^s and N^x inputs, respectively.

Income Flows and Transaction Timing. In the network, each of the N goods sold is assigned a position within a payment chain. Index each good by $j \in \{1, 2, \ldots, N\}$ and let $u_j \in \{0, 1, 2, \ldots\}$ denote the position of good j in its corresponding chain: position 0 if the good is sold spot, 1 if the good is sold to a chained order in position 1, 2 if the good is sold to a chained order in position 2, etc. Since all goods are sold at one unit of numeraire, the price is $1/N$. Hence, total income equals the endowment:

$$y = \frac{1}{N} \sum_{n=0}^{\infty} \sum_{j=1}^{N} \mathbb{I}[u_j = n]. \tag{4}$$

Similarly, each chained order is associated with a payment time. Enumerate each chained order by $i \in \{1, 2, \ldots, N^x\}$ and let $o_i \in \{1, 2, 3, \ldots\}$ denote the position at which chained expenditure i is executed. The final goods obtained from a chained order i is $(1 - \delta)\delta^{o_i}$ per unit, so the final goods obtained from all chained orders is:

$$X = \frac{1}{N} \sum_{i=1}^{N^x} \delta^{o_i}, \tag{5}$$

where X is random, as is o_i. Summing across positions, total chained expenditures are:

$$E^x = \frac{1}{N} \sum_{n=1}^{\infty} \sum_{i=1}^{N^x} \mathbb{I}[o_i = n]. \tag{6}$$

From Payments-Chain Network Budgets to Period Budgets. Wealth evolves within each period as transactions are executed at different moments $\tau(n) \in [0,1]$ within a given period t. Let $W_{\tau(n)}$ denote wealth immediately after all transactions at position n have been completed. Spot expenditures and spot sales occur at $\tau(0)$, so that

$$W_{\tau(0)} = W + \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}[u_j = 0] - E^s. \tag{7}$$

This budget adds the income from goods sold in spot transactions and subtracts spot expenditures from the initial wealth W, yielding $W_{\tau(0)}$. At each later position n, the agent receives income from goods sold and pays for executed chained orders in a given position n:

$$W_{\tau(n)} = W_{\tau(n-1)} + \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}[u_j = n] - \frac{1}{N} \sum_{i=1}^{N^x} \mathbb{I}[o_i = n], \quad \forall n \geq 1. \tag{8}$$

The first term is the carried-over wealth, the second term is the income from sales, and the third term is the payments for chained orders executed at position n. Adding all sequential budgets:

$$W_{\tau(\infty)} = W + \underbrace{\frac{1}{N} \sum_{n=0}^{\infty} \sum_{j=1}^{N} \mathbb{I}[u_j = n]}_{=y} - E^s - \underbrace{\frac{1}{N} \sum_{n=1}^{\infty} \sum_{i=1}^{N^x} \mathbb{I}[o_i = n]}_{=E^x} \Leftrightarrow \frac{W'}{R} + E^s + E^x = y + W. \tag{9}$$

Thus, we obtain an intertemporal budget constraint where $W'/R = W_{\tau(\infty)}$ represents carried over wealth to the next period at gross interest rate R. This budget constraint (9) encodes the evolution of transactions within the period. Importantly, an agent placing chained orders would face uncertainty about the final goods X obtained from these orders. However, as the number of orders becomes large ($N \to \infty$), positions become predictable. Taking the limit as $N \to \infty$, while holding total expenditures fixed, we obtain:

$$\lim_{N\to\infty} X = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N^x} \delta^{o_i} = E^x \sum_{n=1}^{\infty} y^n G(n) = \mathcal{A}(\mu)E^x, \tag{10}$$

where $\mathcal{A}(\mu)$ is the aggregate productivity factor and μ is the economy-wide chained order share. Defining $q(\mu) \equiv \mathcal{A}(\mu)^{-1}$, the budget (9) becomes:

$$\frac{W'}{R} + S + q(\mu)X = y + W, \tag{11}$$

where $S \equiv E^s$ are spot goods obtained via spot orders (final goods equal input goods when orders are bought on the spot), and X are final goods bought via chained orders.

The large transactions limit $N \to \infty$ makes the problem tractable, as budget constraints become deterministic. The cost of placing chained orders, reflected in reduced final goods received, is captured by the average price $q(\mu) \equiv \mathcal{A}(\mu)^{-1} > 1$. Crucially, the specific assignment rule determines how agents interpret this price and budget.

Under this aggregation, agents place orders without knowing when goods will be received. They therefore compute the average value $\mathcal{A}(\mu)E^x$ from chained orders. This formulation implicitly relies on a "big family" assumption in the tradition of Lucas and Stokey (1987), where individual transaction risk is diversified away. The sequential budget constraints also clarify that income from spot transactions may not be pooled: some income inflows may remain idle even when within the same family, and there are pending chained orders. If we want to consider the pooling of funds, the pooled assignment can be used $q = \mathcal{A}^{pool}(\mu)^{-1}$ instead of the random assignment.

Pecuniary Externality. The aggregation of budget constraints also reveals that, in addition to possible inefficiencies from the assignment of orders, the environment features a pecuniary externality when agents take $q(\mu)$ as given. Indeed, $q(\mu)$ is an average and not a marginal cost of obtaining goods via chained orders. This cost depends on μ, affected by individual decisions but not internalized.

As anticipated, one way to eliminate the externality is to centralize transactions through an intermediary that aggregates all orders, treating the timing of transactions, $\tau(n)$, and the output function, $y(n)$, as if these were part of a production technology. That intermediary would treat orders—and their corresponding inputs—as the inputs

$\{E^s, E^x\}$ of an aggregate production function:

$$\mathcal{F}(E^s, E^x) \equiv \mathcal{Y} \times (E^s + E^x) = E^s + E^x \mathcal{A}\left(\frac{E^x}{E^s + E^x}, \mathcal{P}\right).$$

In that case, the intermediary would re-establish efficient pricing by setting $\mathcal{F}_{E^s}$ and $\mathcal{F}_{E^x}$ as the price of spot and chained orders, respectively.

Of course, how realistic it is to assume centralization of payments to establish efficiency is in the eye of the beholder. For that reason, in the business cycle setting below, we contrast an economy with the pecuniary externality with a planner version that avoids it.

Discussion: Reduction of Economic Complexity. We abstract from more realistic features for tractability, enabling the isolation of core mechanisms. Real economies exhibit greater complexity: transactions vary in size (100 vs 200), production networks involve multilateral relationships, and payment timing is more nuanced.[39] The random assignment captures this inherent complexity, suggesting that statistical approaches may be more realistic than the assumption of perfect optimization. Indeed, the pooled and efficient assignments represent increasing levels of coordination: from simple income pooling to complete payment-network optimization. The goal so far is to provide a simple framework to speak to these issues. The next section leverages the aggregation to analyze agents' endogenous decisions between placing spot and chained orders.

4 Endogenous Expenditure Decisions

Next, we develop two dynamic decision problems in which agents confront savings and expenditure-type decisions. The rest of the analysis does not rely on any specific choice of $\mathcal{P}$. However, in all of the illustrations, $\mathcal{P}$ follows the random assignment/

4.1 Consumption-Saving Decisions

We first study the (perfect foresight) problem of entrepreneurs who buy and sell goods in the payments-chain network. Time is discrete $t = 0, 1, 2, \ldots$. At the start of each period, entrepreneurs receive a unit endowment of non-storable inputs available for sale

[39] When transactions differ in size, an agent receiving \$100 but owing \$200 faces complex cash flow problems that require sophisticated optimization—challenges that even large firms struggle with, analogous to NP-hard problems in logistics. We anticipate that methods from statistical physics could be used to address the combinatorial challenges arising from transactions of different sizes.

and purchase inputs from others that are converted into final consumption goods. Entrepreneurs have log-utility over final consumption goods. At the start of each period, entrepreneurs place and receive spot and chained orders, as above. For this section, they take the sequence of prices $\{q_t\}_{t\geq 0}$, and interest rates, $\{R_{t+1}\}_{t\geq 0}$, as given.

At the start of t, the entrepreneur chooses *spot* purchases $S_t \geq 0$ (requiring immediate funds) and *chained* goods $X_t \geq 0$ (purchased via chained orders). Their sum, $S_t + X_t$, yields period utility, indicating that final goods are perfect substitutes. Total expenditures, $E_t = S_t + q_t X_t$, are as in the aggregated budget constraint from the previous section. Recall that we treat inputs as a numeraire while $q = \mathcal{A}^{-1}$ translates chained expenditures into goods. The entrepreneur enters period t with debt B_t (corresponding to negative wealth in the previous section) and carries over B_{t+1} at rate R_{t+1} to next-period debt.

We incorporate two constraints that make this problem meaningful. First, a natural borrowing limit, $B_{t+1} \leq \bar{B} \equiv 1/(1-\beta)$. Second, spot purchases require liquidity. A *spot borrowing line* (SBL) $\tilde{B}_t$ limits intraperiod liquidity:

$$S_t \leq \bar{S}_t \equiv \max\{\tilde{B}_t - B_t,\, 0\}.$$

The entrepreneur spot and chained expenditures to maximize lifetime $\sum_t \beta^t \log(S_t + X_t)$. Given total expenditures E_t, since $q_t \geq 1$, by cost minimization, the entrepreneur makes as many spot expenditures as the SBL permits. Thus, given E_t, spot and chained expenditures are:

$$S_t = \min\left\{\bar{S}_t, E_t\right\}, \qquad X_t = \left(E_t - \min\left\{\bar{S}_t, E_t\right\}\right)/q_t. \tag{12}$$

Given this optimal expenditure split, the entrepreneur's intertemporal utility maximization can be cast into a dynamic programming problem:

Problem 1. (Entrepreneur's Recursive Problem): *Given B_0 and $\left\{\tilde{B}_t, R_{t+1}, q_t\right\}_{t\geq 0}$, the entrepreneur chooses debt holdings, $\{B_{t+1}\}_{t\geq 0}$, that solve:*

$$V_t(B) = \max_{B' \leq \bar{B}} \log(S + X) + \beta V_{t+1}(B')$$

where S and X are given by (12) *and total expenditures by $E = B' R_{t+1}^{-1} + 1 - B$.*

The Bellman equation expresses the entrepreneur's problem as a choice over intertemporal debt. The choice of carried-over debt, B', determines the period's expenditures, through (9), and the optimal split (12) yields S and X given the level of expenditures.[40] When short-term means of payment are scarce, agents resort to chained orders, understanding their higher cost.

The spot borrowing line (SBL), $\bar{S}_t$, represents short-term funds sourced from credit lines or trade credit, as explained earlier. These loans carry interest only if rolled over into future debt. Of course, if agents could borrow further intra-period spot funds beyond $\bar{S}$, they would as long as the intra-period loan is less than $1/q_t$. The SBL differs fundamentally from the hard debt limit, $\bar{B}$, limiting intertemporal borrowing B_{t+1}. Entrepreneurs can roll over SBL usage into future debt B_{t+1} without immediate principal repayment. This is akin to a credit card with a limit: beginning-of-period expenditures are limited by the outstanding balance, which, in turn, carries over interest on the end-of-period balance.[41]

Characterization. We identify two debt thresholds that determine whether entrepreneurs can purchase all of their expenditures through spot orders or, conversely, whether they can make any spot expenditures at all.

Lemma 1 (Expenditure Threshold Points)**.** *$S_{t+1} = 0$ if and only if $B_{t+1} > \tilde{B}_{t+1}$. Define the efficiency threshold, $B^{\star}_{t+1} \equiv R_{t+1}(\tilde{B}_t - 1)$. Then, $X_t > 0$ if and only if $B_{t+1} > B^{\star}_{t+1}$.*

The following period's SBL $\tilde{B}_{t+1}$ determines whether entrepreneurs can make any spot purchases in the subsequent period. In turn, when carried-over debt exceeds the efficiency threshold $B^{\star}_{t+1}$, this indicates spending beyond the SBL and, therefore, the use of chained expenditures.

These thresholds are critical for the characterization because debt levels affect the availability of spot expenditures and, thus, the price at which goods are bought at the margin, i.e., *marginal prices*, relevant for marginal decisions. The *marginal expenditure price*, $\tilde{q}^E_t(B') \equiv 1 + (q_t - 1)\mathbb{I}_{[B' \geq B^{\star}_{t+1}]}$ indicates the marginal cost of final goods when carried-over debt is B'. The *marginal borrowing price* is $\tilde{q}^B_{t+1}(B') \equiv 1 + (q_{t+1} - 1)\mathbb{I}_{[B' > \tilde{B}_{t+1}]}$ indicates the marginal cost of a good tomorrow purchased with marginal savings today.[42] Marginal prices capture the consumption-saving tradeoffs that determine the entrepreneur's optimal debt, via a generalized Euler equation:

---

[40] The time index in the value function reflects the evolution of $\{R_{t+1}, \tilde{B}_t, q_t\}$.

[41] Financial intermediaries often set different limits as a way to tighten short-term credit without unfeasibly demanding principal repayment, which could trigger costly defaults for both parties.

[42] If carried over debt exceeds the efficiency threshold, reducing debt requires sacrificing $1/\tilde{q}^E_t$ units

Proposition 2 (Entrepreneur's First-Order Condition)**.** *Consider an increasing sequence of $\tilde{B}_t$. Any solution $\{B_{t+1}\}_{t\geq 0}$ to the entrepreneur's problem satisfies:*

$$\frac{E_{t+1}}{\beta E_t}\frac{Q_t}{Q_{t+1}} = \frac{R_{t+1}}{\tilde{q}^B_{t+1}(B_{t+1})/\tilde{q}^E_t(B_{t+1})} \text{ if } B_{t+1} \neq B^\star_{t+1}, \quad \beta q_t R_{t+1} \geq \frac{E_{t+1}}{E_t} \geq \beta R_{t+1} \text{ if } B_{t+1} = B^\star_{t+1},$$

where$Q_t \equiv E_t/(S_t + X_t)$ is the average prices of final goods.

Despite its unconventional format, this Euler equation retains a familiar interpretation: at interior solutions ($B_{t+1} \neq B^\star_{t+1}$), marginal rates of substitution across periods equal the return on savings corrected for price differences between spot and chained consumption. The left side represents the ratio of marginal rates of substitution (expressed as the ratio of expenditures over average prices yielding the ratio of consumption consistent with log utility), while the right side captures returns adjusted by "inflation" in terms of marginal prices. In turn, the inequalities follow when future debt is the efficient debt level, a corner solution.

This Euler condition is necessary but not sufficient, as the value function is not concave: multiple sequences of future debt may satisfy the condition starting from an initial state. This non-convexity opens the door to debt-overhang leading to hysteresis in general equilibrium.

Stationary Dynamics. We now characterize the stationary dynamics of the entrepreneur's problem, i.e., when exogenous variables he are constant, $\tilde{B}_t = \tilde{B}_{ss}$, $q_t = q_{ss}$, and $R_t = R_{ss} = 1/\beta$. We focus on $\tilde{B}_{ss} > 1$ and define $B^\star_{ss} \equiv B^\star(1/\beta, \tilde{B}_{ss})$.

Proposition 3 (Stationary Entrepreneur Problem)**.** *In a stationary solution: If $B_0 \leq B^\star_{ss}$, then $B_t = B_0, \forall t$. Moreover, there exists $B^h > \tilde{B}_{ss}$ such that: If $B_0 \in (B^\star_{ss}, B^h)$, then $B_t \to B^\star_{ss}$ in a a finite number of periods. If $B_0 > B^h$, then $B_t = B_0, \forall t$.*[43]

Figure 7 brings together the theoretical insights of the proposition. To illustrate the forces at play, we overlay the entrepreneur's value function with two benchmark cases: $\bar{V}$, the value function when the spot borrowing limit is unbounded, ($\tilde{B}_t = \infty$) and the entrepreneur can finance any desired of spot expenditures, and, $\underline{V}$, where no liquidity is available ($\tilde{B}_t = 0$) and, hence, no spot expenditures are possible.

---

of consumption. If future debt exceeds the spot borrowing line, savings yield $1/\tilde{q}^B_{t+1}$ goods per unit of savings.

[43]Appendix D.1 provides a complete characterization of B^h and the number of periods needed to arrive at $B^\star_{ss}$. At B^h there is indifference between deleveraging or not.

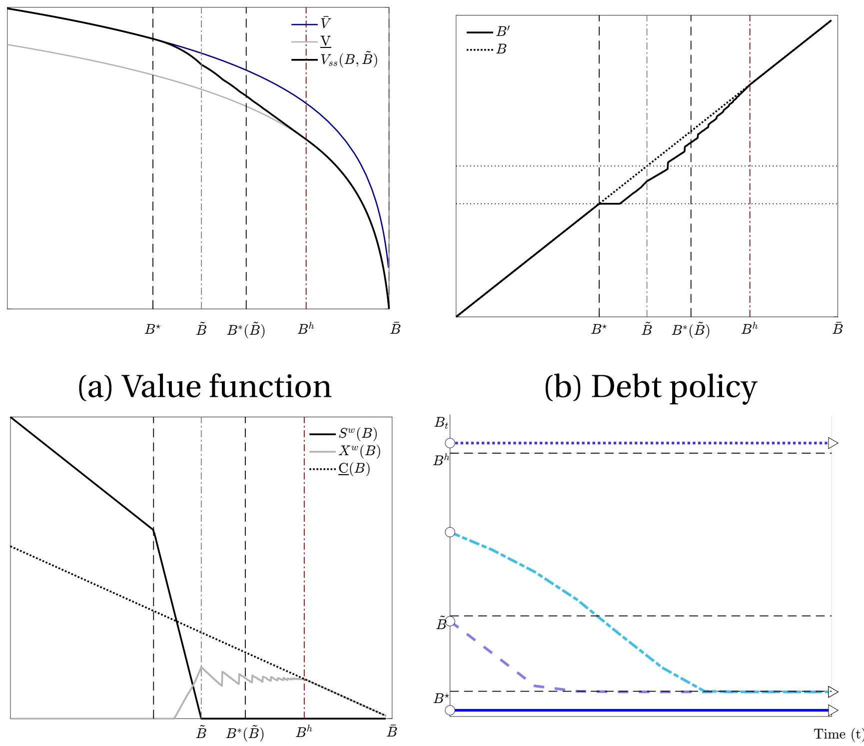
Panel (a) demonstrates how the actual value function deviates from both perfect liquidity ($\bar{V}$) and zero liquidity ($\underline{V}$) benchmarks for intermediate debt levels. For debt levels below $B^{\star}$, the value function V coincides with $\bar{V}$: entrepreneurs are unconstrained and consume the annuity value of their endowment net of interest payments entirely via spot purchases, showcasing that for sufficiently low debt levels, the SBL is irrelevant. When debt exceeds the high threshold B^{h}, the value function lies on $\underline{V}$, indicating that entrepreneurs consume exclusively via chained purchases forever. Crucially, even when highly indebted, the entrepreneur could potentially save to access better prices in the future. However, he chooses not to because the benefits of deleveraging come too far into the future.

Consistent with Panel (a), Panel (b) shows non-monotonicity in the optimal debt policy function. Entrepreneurs with debt below $B^{\star}_{ss}$ have constant debt levels; entrepreneurs with debt above B^{h} remain permanently at their initial debt levels. Only those with moderate debt in $(B^{\star}_{ss}, B^{h})$ gradually deleverage toward the efficient threshold $B^{\star}$ by choosing future debt (solid line) below current debt (dashed line).

Panel (c) illustrates how consumption policies differ across debt levels. When debt is below (above) $B^{\star}_{ss}$ (B^{h}), the entrepreneur consumes the annuity of his wealth as spot (chained) consumption. For debt levels slightly below B^{h}, we observe a reduction in overall consumption to deleverage. The chain-saw pattern of the optimal policy is a feature that reflects a trade-off between smoothing consumption and shortening the number of periods to arrive at the efficient debt level. Panel (d) traces the debt dynamics that emerge from the non-monotonic policy functions: Starting from different initial debt levels, entrepreneurs either remain at their initial debt levels or temporarily delever toward the efficient debt threshold.

The main takeaway is that, unlike canonical financial frictions models, entrepreneurs lack incentives to delever, even when their discount factor equals the real interest rate. Indeed, if sufficiently indebted, they prefer to stay in debt and pay the high price of chained orders. This distinction occurs as the benefits of deleveraging do not happen at the margin. On the contrary, the partitioning of the state space into two regions with different domains of attraction is a shared feature with models where, due to financial constraints, there are rewards to bringing debt to a certain level. Whereas the rewards here stem from accessing better prices, in other models rewards stem from avoiding roll-over risk (Cole and Kehoe, 2000; Aguiar et al., 2015) or technologies requiring scale

(Buera and Shin, 2013).[44] The lack of incentives to delever has significant consequences for the general equilibrium that are aggravated by the pecuniary externalities.



(a) Value function

(b) Debt policy

(c) Consumption policy

(d) Transitions

Figure 7: Bellman Equation and Policy Functions: Entrepreneur's Problem

Note: Figures computed using value function iteration with $\beta = 0.8$, $R = \beta^{-1}$, $q = 1.75$, and $\tilde{B} = 0.4 \cdot \bar{B}$.

4.2 Investment Decisions

We can also study how payment chains can affect investment and production by incorporating these features into the entrepreneur's problem. For that, we study a problem where both consumption and investment must be purchased in the payment network and part of yesterday's purchases are inputs for today's production.

Problem 2. (Entrepreneur's Problem with Production): *Given* B_0 *and* $\left\{\tilde{B}_t, R_{t+1}, q_t\right\}_{t\geq 0}$,

[44] A continuous-time limit akin to those studied in roll-over risk models shows that the saw-like pattern of numerical solution disappears and that the value functions are isomorphic—the saw like pattern appears in discrete time due to the integer periods of arrival to unconstrained debt position; see the online note, "Consumption–Saving Problem with Wealth Rewards."

the entrepreneur chooses debt-holdings and investment $\{B_{t+1}, K_{t+1}\}_{t\geq 0}$, *the solution to:*

$$V_t(B,K) = \max_{B'\leq\bar{B},K'\geq 0} \log(S+X-K') + \beta V_{t+1}(B',K')$$

where S *and* X *are given by* (12) *and total expenditures by* $E = B'R_{t+1}^{-1} + f(K) - B$.

Instead of an endowment, the entrepreneur has $f(K)$ goods produced with input (e.g., capital or inventory) K acquired in the previous period. In turn, currently purchased goods, $S+X$, are distributed between consumption, C, and inputs acquired for next-period production, K', assuming full depreciation to keep the analysis as simple as possible.

Recall that in the endowment problem above, the entrepreneur's stationary problem reaches two types of steady-states: either all-spot-expenditures or all-chained-expenditures steady states. The same two types of steady states are reached in the problem with investment. Furthermore, away from steady state, the same incentive to reduce debt and to confront better prices for moderate debt levels is present in this problem, leading to types of steady states and an intermediate region, a domain of attraction toward the efficient steady state.

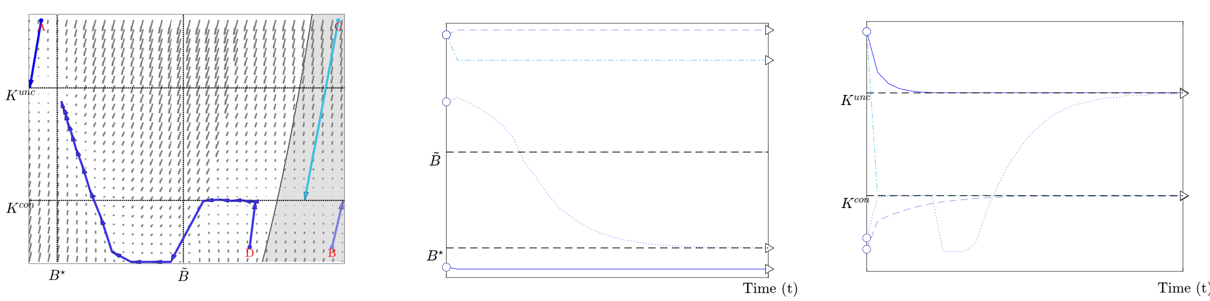
The different insight emerging from the version with production is that the higher cost of chained goods also distorts the optimal investment scale. If the entrepreneur only makes spot purchases, the return on investment equals the rate of interest: $f'(K_{ss}) = R = \frac{1}{\beta}$. However, when all expenditures are chained, the first-order condition for inputs becomes: $f'(K_{ss}) = qR = \frac{q}{\beta}$. Hence, the marginal product of capital exceeds the rate of return on saving and investment, as in other distorted economies. In intermediate cases, investment is distorted by average prices.

Figure 8 illustrates the domains of attraction toward the two classes of steady states and the rich dynamics that emerge with production. Panel (a) shows the phase portrait[45] with arrows indicating the direction of change for both state variables (B,K) of the model's solution together with some trajectories. The shaded region represents the domain of attraction toward the disrupted steady state with only chained expenditures and inefficient investment. Four representative paths demonstrate the range of possible dynamics. Path A begins with low debt and excess capital: convergence to the efficient steady state quickly as excess income is used to reduce debt, reaching equilibrium

[45] A phase portrait draws the discrete vector field—where arrows indicate the direction and magnitude of change from each point—one (B,K) to (B',K').

with undistorted investment while maintaining the pattern of only spot expenditures. Paths B and C start with high initial debt but different capital levels—both converging to the steady state with low investment and only chained expenditures. The most interesting case, Path D, starts with intermediate debt and low capital. Initially, the entrepreneur accumulates both debt to increase capital and reach an efficient scale. It begins to payoff debt and once debt falls sufficiently to enable some spot purchases, investment is reduced to accelerate the deleveraging until the efficient level of capital is reached and the entrepreneur makes only spot expenditures.

Panels (b) and (c) trace the corresponding debt and capital paths over time, confirming that the same two-steady-state pattern holds with production: entrepreneurs either converge quickly to efficiency or remain trapped in a state of persistent underinvestment. The same pattern of debt overhang, when debt starts high but deleveraging when debt is low, is observed in panel (b). Panel (c) shows how capital evolves depending on the marginal prices.



(a) Phase Portrait (b) Path of Debt (c) Path of Capital

Figure 8: Investment and Leverage Decisions

Note: Figures are calculated using value function iteration with $\beta = 0.8$, production function $f(k) = k^\alpha$ with $\alpha = 0.5$, $q = 1.25$, $R = \beta^{-1}$, and $\tilde{B} = 0.2 \cdot \bar{B}$. Panel (a) corresponds to the phase-space diagram. Panels (b) and (c) correspond to the transitions for debt and capital respectively.

5 General Equilibrium

The previous section analyzes how entrepreneurs choose between spot and chained purchases given exogenous terms. We now embed these decisions into a general equilibrium where prices and interest rates are endogenous revealing how individual payment choices create aggregate externalities and providing insights about crises.

Environment. We provide a simple setting with endogenous savings and expenditures decisions. Timing remains identical to the partial equilibrium problems above: payments are executed within the period and debt is carried over across periods. As in Kiyotaki and Moore (1997b), the economy consists of two groups of agents. *Entrepreneurs* are the same agents studied in the previous section, but we specialize their aggregate input endowment to $y = 1$. The new agents are *creditors* who hold positive financial wealth but receive no endowment. Both share the same log preference. To consume, creditors always carry positive wealth, D_t, earning the gross return R_{t+1} between periods. Creditors serve as the economy's savers, providing the funds that entrepreneurs borrow.[46]

Problem 3 (Creditor's (Recursive) Problem)**.** *Given D_0 and $\{R_{t+1}\}_{t\geq 0}$, the creditor chooses deposit holdings $\{D_{t+1}\}_{t\geq 0}$ to solve:*

$$V_t(D) = \max_{D' \geq 0} \log(C) + \beta V_{t+1}(D')$$

subject to the intertemporal budget constraint: $C + R_{t+1}^{-1} D' = D$.

The solution is standard to logarithmic preferences: $C_t = (1-\beta)D_t$ and $D_{t+1} = \beta R_{t+1} D_t$ for all $t \geq 0$. Their total expenditure each period is $C_t = (1-\beta)D_t$. Implicitly, we use that creditors always have positive wealth, as they only make spot expenditures.[47]

Equilibrium. Since entrepreneurs supply the input endowment, $y = 1$, which is the numeraire, *clearing in the input market* requires total expenditures to equal the input endowment, $E_t^s + E_t^x = 1$. *Clearing in the asset market* requires the creditors' deposits exactly equal entrepreneurs' debt: $D_t = B_t$. Spot expenditures are the sum of the creditors' consumption (who only make spot purchases) and entrepreneurs' spot purchases: $E_t^s = C_t + S_t$. Chained expenditures come only from entrepreneurs: $E_t^x = q_t X_t$. From Proposition 1, chained expenditures are $\mu_t = q_t X_t$, which determines *a price consistency condition* $q_t = \mathcal{A}^{-1}(\mu_t)$. Because spot expenditures translate one for one to output we have that *final goods consumption equals total production* $C_t + S_t + X_t = \mathcal{Y}(q_t X_t)$.

Definition 2 (Symmetric Competitive Equilibrium)**.** *Given a sequence of borrowing limits $\{\tilde{B}_t\}_{t\geq 0}$, sequence $\{B_t, C_t, S_t, X_t, R_t, q_t\}_{t\geq 0}$ is a **symmetric competitive equilibrium** it*

[46] Creditors play a crucial role: they close the asset market and their spot purchases guarantee there's a mass of spot expenditures that prevents the entire economy from collapsing if the entrepreneurs' debt exceeds the SBL.

[47] That is, their budget constraint also takes the form of (11) with $X = 0$.

satisfies $\forall t$*:* ***I. Optimization*** *(given* $\{R_t, q_t\}_{t\geq 0}$*, the sequences* $\{D_t, B_t, S_t, X_t\}_{t\geq 0}$ *solve the entrepreneur's problem and* $\{D_t, C_t\}_{t\geq 0}$ *the creditor's problem),* ***II. Asset and input markets clear, III. Final goods consumption equals production****, and* ***IV. Price consistency****.*

The equilibrium definition is standard, except that q_t and $\mathcal{Y}_t$ depend on the amount of chained expenditures. Despite the richness of network features, the economy operates as an economy with production externalities on $\mathcal{Y}$ and pecuniary externalities on q, encoding the payments technology and delays in the payments-chain network.

Equilibrium Dynamics: Recursive Representation. A virtue of this setting is that it renders a tractable way to characterize the equilibrium recursively. Given that $D_t = B_t$, B_t is the single endogenous state. The state variable is the vector $\{B_t, \tilde{B}_t, \tilde{B}_{t+1}\}$. A recursive equilibrium is a mapping from the state to the vector of prices R_{t+1}, q_t, and allocations, $\{C_t, S_t, X_t\}$, and future endogenous state B_{t+1}. Because a recursive representation is possible, we drop time subscrits adopting standard recursive notation. Since only B is endogenous, we seek a mapping $\mathcal{B}$ that determines tomorrow's debt, B', as a function of today's state: $B' = \mathcal{B}(B, \tilde{B}, \tilde{B}')$.

We begin by establishing that the time t allocation is a only function of B and $\tilde{B}$.

Proposition 4 (Equilibrium Allocation given $\{B, \tilde{B}\}$)**.** *Given debt* B *and spot borrowing limit* $\tilde{B}$*, the equilibrium allocation is characterized by:*

1. ***Creditor consumption****:* $C(B) = (1-\beta)B$ *(creditor log expenditures)*
2. ***Entrepreneur expenditures****:* $E(B) = 1 - (1-\beta)B$ *(input-market clearing)*
3. ***Spot purchases****:* $S(B, \tilde{B}) = \min\{\max\{0, \tilde{B} - B\}, E(B)\}$ *(spot maximization)*
4. ***Chained expenditures****:* $\mu(B, \tilde{B}) = E(B) - S(B, \tilde{B})$ *(definition of total expenditures)*
5. ***Network price****:* $q(B, \tilde{B}) = \mathcal{A}^{-1}(\mu(B, \tilde{B}))$ *(definition)*
6. ***Chained goods****:* $X(B, \tilde{B}) = \mu(B, \tilde{B})/q(B, \tilde{B})$ *(definition)*
7. ***Average price****:* $Q(B, \tilde{B}) = E(B)/(S(B, \tilde{B}) + X(B, \tilde{B}))$ *(definition).*

This proposition shows that the current debt level and SBL yield the entire allocation of goods. What remais is to characterize the evolution of debt, i.e., finding $\mathcal{B}$. This debt mapping emerges from combining the entrepreneur's and creditor's Euler equations. From the creditor's Euler equation, the gross real interest rate can be expressed as a function of B and B'; namely, $R(B, B') = B'/(\beta B)$. Thus, we can replace this gross rate in the entrepreneur's interior Euler equation, Proposition 2, to obtain

$$\underbrace{\frac{B}{E(B)} \cdot Q(B, \tilde{B})}_{\equiv \mathcal{E}(B;\tilde{B})} = \underbrace{\frac{B'}{E(B')} \cdot Q(B', \tilde{B}') \cdot \frac{q^E(B'; B, \tilde{B}, \tilde{B}')}{q^B(B'; \tilde{B}')}}_{\equiv \mathcal{E}'(B';B,\tilde{B},\tilde{B}')} \tag{13}$$

where $q^E(B'; B, \tilde{B}, \tilde{B}') \equiv 1 + (q(B, \tilde{B}) - 1)\mathbb{I}_{[B' \geq B^\star(\tilde{B})]}$ and $q^B(B'; \tilde{B}') \equiv 1 + (q(B', \tilde{B}') - 1)\mathbb{I}_{[B' > \tilde{B}']}$ are the marginal expenditure and borrowing prices as functions of the state.

When interior solutions to the entrepreneur's problem are are valid, the debt mapping is determined by solving $\mathcal{E}(B; \tilde{B}) = \mathcal{E}'(B'; B, \tilde{B}, \tilde{B}')$ for B'. Due to the discontinuous nature of marginal prices, this equation may have multiple solutions, as in partial equilibrium. However, the partial equilibrium problem tells us how to select the appropriate solution, so long as the entrepreneur wants to delever at given prices:

$$\mathcal{B}(B, \tilde{B}, \tilde{B}') = \max\left\{B^\star(\tilde{B}), \arg\min_{B'}\{\mathcal{E}(B; \tilde{B}) = \mathcal{E}'(B'; B, \tilde{B}, \tilde{B}')\}\right\}.$$

This mapping ensures that entrepreneurs never choose debt levels below the threshold $B^\star(\tilde{B})$ where they would be unconstrained, and among multiple solutions, they select the one that maintains continuity in their value function. It is a non-linear implicit difference equation provides a recipe for solving the model: first, we solve for the equilibrium allocation and price functions as a function of $\{B, \tilde{B}\}$. Then, we implicitly solve for $\mathcal{B}$. We exploit this feature to obtain a graphical device to analyze equilibria and obtain several properties.

5.1 Implication I: Transitional Dynamics and Hysteresis

Steady States. In a steady state, the SBL, $\tilde{B}_{ss}$, and debt must be constant: $B_{ss} = B$. From the creditor's solution, we know that a necessary and sufficient condition for any steady state is that the real interest rate is constant at $R_{ss} = \beta^{-1}$. Thus, a specific debt level constitutes a steady state if, given this interest rate, and the resulting equilibrium price q, entrepreneurs optimally choose to maintain that same debt level, a condition we studied in partial equilibrium.

Since at a steady state with $R_{ss} = 1/\beta$ entrepreneurs either only make spot or or only make chained expenditures, we conclude there are two types of steady state in general equilibrium—there's a continuum of steady states depending on B_0. In an *undisrupted steady state*, debt remains below $B^{\star}$, entrepreneurs exclusively make spot purchases, $q = 1$, payments are immediate, and the economy operates at full potential with $\mathcal{Y}_{ss} = 1$. In a *disrupted steady state*, entrepreneurs rely entirely on chained purchases, resulting in $q > 1$, delayed payments, and output below potential: $\mathcal{Y}_{ss} < 1$.

Corollary 1 (Steady-State Characterization)**.** *Fix $\tilde{B}_t = \tilde{B}_{ss} > 0$. Then, for all t:*

i. The economy is in an undisrupted steady state when $B_t \leq B^{\star}(\beta^{-1}, \tilde{B}_{ss})$.

ii. The economy is in a disrupted steady state when $B_t \geq B^{h}(\beta^{-1}, \tilde{B}_{ss})$.

The disrupted steady state represents an inefficient hysteresis region where excessive debt permanently impairs the economy's payments efficiency. The economy is constrained inefficient, as we show below, because neither agent internalizes the effects of their expenditures on q.[48]

Domain of Attraction Toward Efficient Steady States. For debt levels between the two steady-state regions, the economy may experience transitional dynamics toward the undisrupted steady-state region. The domain of attraction toward undisrupted steady states depends on whether the aggregate Euler equation $\mathcal{E}(B; \tilde{B}) = \mathcal{E}'(B'; B, \tilde{B}, \tilde{B}')$ has a solution with $B' < B$—that is, entrepreneurs choose to delever. This domain of attraction toward undisrupted steady states has an upper bound,[49] B^{*} given by:

$$B^{*}(\tilde{B}) \equiv \tilde{B}/(S(\tilde{B}, \tilde{B}) + X(\tilde{B}, \tilde{B}) + C(\tilde{B})) \geq \tilde{B}.$$

Corollary 2 (Convergence to Efficiency)**.** *Let $\tilde{B}_t = \tilde{B}_{ss}$. For any $B_0 < B^{*}(\tilde{B}_{ss})$:*

1. *If $B_t \in (B^{\star}(\beta^{-1}, \tilde{B}_{ss}), B^{*}(\tilde{B}_{ss}))$, then $B_{t+1} < B_t$ (deleveraging occurs)*

2. *If $B_t \leq B^{\star}(\beta^{-1}, \tilde{B}_{ss})$, then $B_{t+1} = B_t$ (steady state reached).*

Payments-chain delays are only *temporary* when the SBL is at steady state, provided that debt starts in the domain of attraction toward efficient steady states.

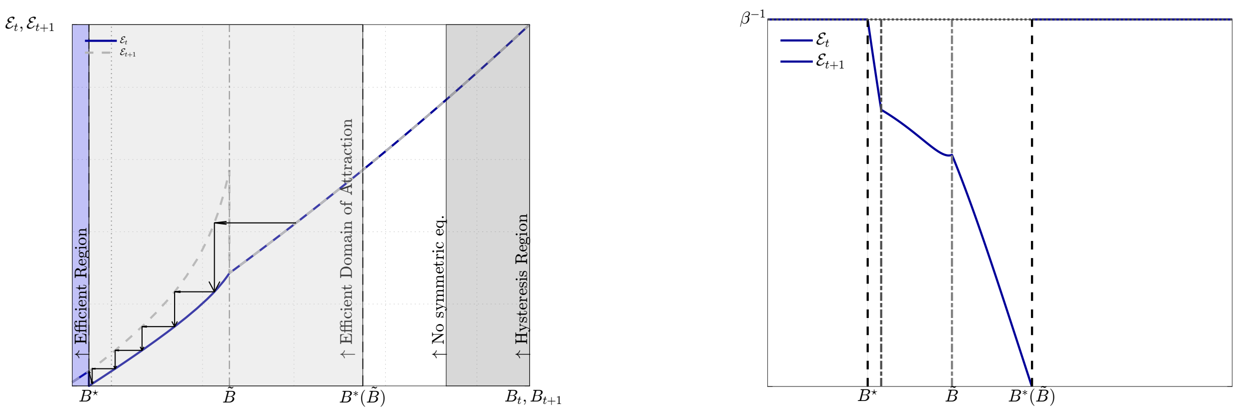
Figure 9 illustrates the dynamics encoded in $\mathcal{B}$. Panel (a) plots the functions $\mathcal{E}$ (solid blue) and $\mathcal{E}'$ (dashed gray) against debt levels, holding $\tilde{B}$ fixed. The arrows trace the

[48]This mechanism provides a theoretical foundation for understanding persistent periods of low productivity among highly indebted firms observed during extended economic stagnation.

[49]For debt levels exceeding this upper bound, the economy does not transition toward the efficient steady state in a symmetric equilibrium.

equilibrium path of debt: when the functions intersect below the 45-degree line, entrepreneurs deleverage ($B' < B$). The convergence region extends from $B^\star$ to B^*, while the gray shaded area represents the hysteresis region. Panel (b) shows how equilibrium interest rates vary across debt levels, remaining at β^{-1} in both steady-state regions.

Finally, we note that the domains of attraction toward efficient steady states and the hysteresis region are disconnected. When debt falls in the intermediate range $\{B^*, B^h\}$, symmetric competitive equilibria do not exist. In this middle region, entrepreneurs individually would want to deleverage, but in general equilibrium, this would require interest rates far below β^{-1}, for which no solutions exist to the equilibrium condition. We leave the analysis of asymmetric equilibria for future work.



(a) Transition plot

(b) Equilibrium rates

Figure 9: Equilibrium Phase Diagram and Hysteresis Region

Note: Figures are calculated using value function iteration with $\beta = \delta = 0.95$, and $\tilde{B} = 0.2 \cdot \bar{B}$. Panel (a) shows the phase diagram of B_t constructed using $\mathcal{E}(B'; B)$ and $\mathcal{E}'(B'; B, \tilde{B}, \tilde{B}')$. Arrow indicate the projection from one point in the Euler equation to another. The projection to the x-axis traces that path of debt. Panel (b) plots $R(B, \tilde{B}, \tilde{B}')$ over $[B^\star - 1,\ B^* + 2]$. The shaded area corresponds to the hysteresis region.

5.2 Implication II: Credit Crunches and a Resilience Paradox

Credit Crunches. We now analyze the economy's response to temporary disruptions in short-term funding, i.e., a temporary tightening of borrowing limits $\{\tilde{B}_t\}$ that goes to zero for some time and then mean-reverts to its steady-state level.[50] As long as initial debt satisfies $B_0 \leq B^*(\tilde{B}_{ss})$ and $\tilde{B}_{ss} > 1$, the economy eventually reverts to an undisrupted steady state.

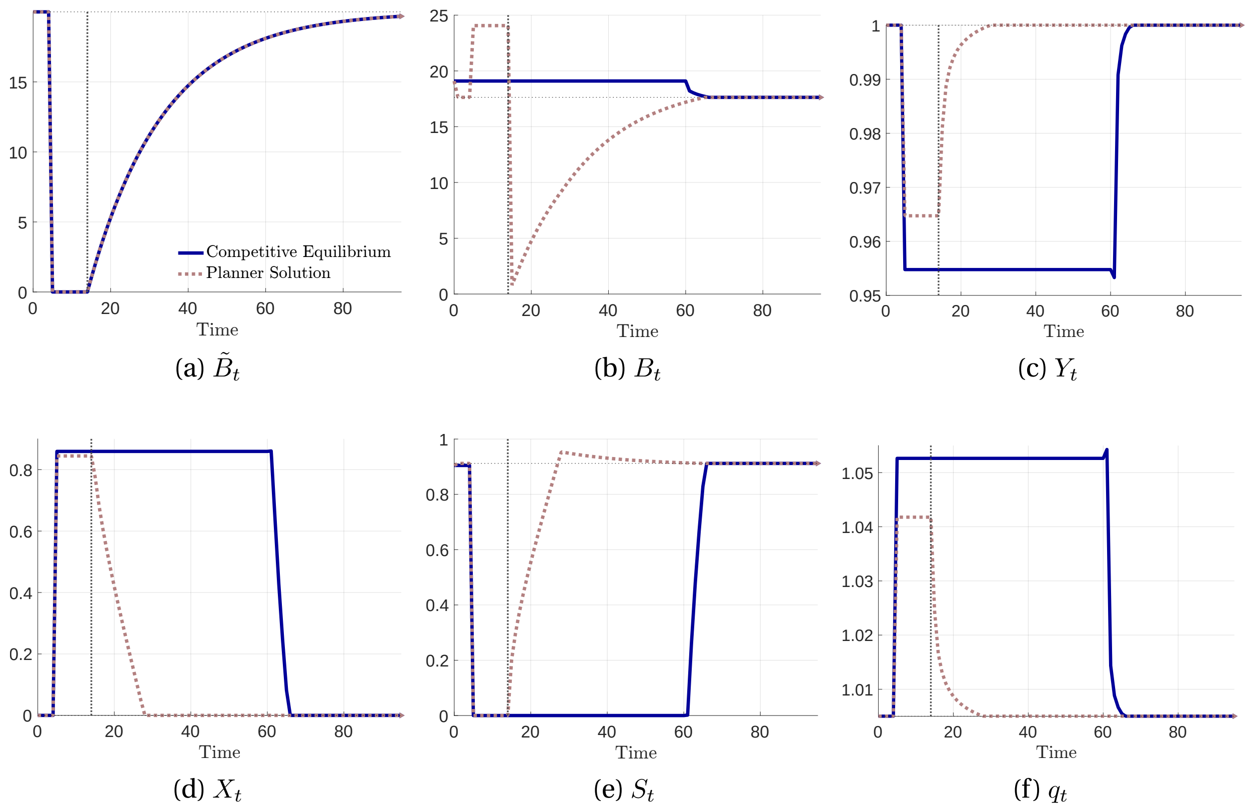
[50]The debt mapping $\mathcal{B}$ allows us to characterize the full transition path by running forward the difference equation for B.

Figure 10 illustrates a transition following a credit crunch starting from an undisrupted steady state. The solid line shows the competitive equilibrium response, the dashed line represents the constrained planner's solution, discussed after. The transition exhibits three distinct phases: The *extreme phase* occurs when $\tilde{B}_t$ falls to zero, eliminating access to short-term credit and no spot expenditures are feasible by the entrepreneur. A *hysteresis phase* where some spot expenditures would be possible if debt was reduced. In this phase, debt remains constant and TFP remains depressed at its lowest value even though the SBL is recovering. Debt levels remain constant during this phase because entrepreneurs keep chained expenditures constant. Finally, there's a *deleveraging phase* where debt is reduced for a while, TFP recovers quickly and full potential is restored.

The most interesting aspect of this transition is the contrast between the hysteresis phase and the deleveraging phase. During the deleveraging phase, entrepreneurs anticipate that prices will improve the following period because for the current lever of B, the recovery of the SBL will allow some spot consumption on the margin. Thus, it is worth sacrificing chained expenditures for spot expenditures. This desire reduces the level of debt, reducing the interest rate, and increases a positive externality which is why output recovers quickly. While this recovery could have been accelerated, the hysteresis region is characterized by depressed TFP and no willingness to delever. On the margin, entrepreneurs understand that they are saving on chained expenditures to make more chained expenditures the subsequent period. Thus, they keep their debt constant. The planner's solution, follows a completely different path.

A Resilience Paradox. A natural question is whether economies with faster payments are always more resilient to credit crunches? The answer is no. Next, we illustrate a resilience paradox: economies with lower δ (greater payment delays) may exhibit superior resilience to credit shocks, despite always experiencing larger efficiency losses for a given μ—recall Figure 6c.

The mechanism underlying this resilience paradox operates through differential deleveraging incentives before a shock. A lower δ means longer payment delays in the payments-chain network, making chained purchases more costly relative to spot purchases. Thus, greater costs create stronger incentives for entrepreneurs to delever. Therefore, a lower δ induces a fundamental trade-off: greater payment inefficiency in the short run, given a level of chained expenditures, but reduced vulnerability to credit



(a) $\tilde{B}_t$ (b) B_t (c) Y_t

(d) X_t (e) S_t (f) q_t

Figure 10: Efficient and Competitive Transition after a Credit Crunch

Note: This figure reports a numerical example of a credit-crunch episode. Figures are calculated using value function iteration with $\beta = 0.995 = \delta = 0.995$, and $\tilde{B}_{ss} = 0.1 \cdot \bar{B}$. The figure is computed by setting 5 periods at an initial steady state, followed by a credit crunch where $\tilde{B}_t$ is set to zero for 10 periods. From $t = 15$, $\tilde{B}_t = 0$ on, $\tilde{B}_t$ returns to steady state according to an AR(1) process with coefficient 0.95. The planner solution uses the Pareto weights consistent with the terminal steady state, $\tilde{B}_t = \tilde{B}_{ss}$.

crunches through the incentives delever.

Figure 11 demonstrates the paradox through a calibrated comparison. We examine two economies that begin with identical debt levels but differ in their payment efficiency parameter δ. After allowing sufficient time for each economy to reach its preferred debt level, we subject both to an identical credit shock—the reduction in $\tilde{B}_t$ shown in Panel (a).[51]

Panel (b) reveals the key mechanism: when the shock hits, the low-δ economy (solid line) has achieved significantly lower debt levels than the high-δ economy (dashed line) due to stronger deleveraging incentives from costlier chained purchases. Panel (c) shows the paradox: despite being less efficient at processing payments, the low-δ economy experiences a smaller output decline during the credit crunch. The the credit crunch occurs at a time when debt is lower in the more inefficient economy, which

[51]The shock is anticipated, but the same is true if it is not.

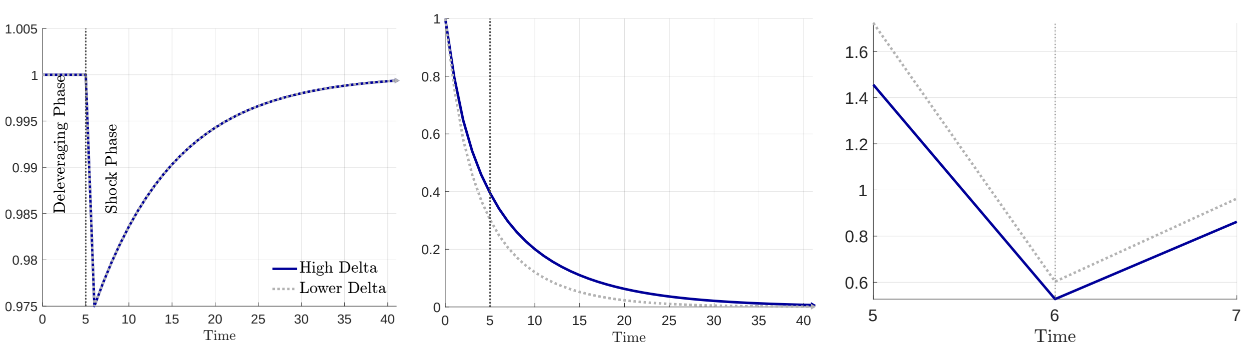
reduced debt faster to avoid the reliance on more expensive chained purchases.

The resilience paradox requires some qualifications. (1) The paradox only manifests when both economies are in a transition toward an efficient steady-state and we let enough time so that the low-δ exhibits less debt. The low-δ deleverages faster, but needs time to become more resilient. Indeed, if the shock happens and both economies have the same level of debt, the low-δ economy will always do worse. (2) For extremely severe shocks, e.g., $\tilde{B}_t = 0$, the low-δ economy will necessarily perform worse.

Despite the qualifications, the resilience paradox reveals an economic insight: payment system inefficiencies, while costly during normal times, can induce deleveraging that builds shock resilience. In fact, the empirical analysis is suggestive of this deleveraging-to-avoid-costs pattern: sectors with the slowest transaction technology rely less on chained transactions (Panel (b) of Figure 2). The takeaway is that economies with more efficient payment systems may inefficiently maintain higher debt levels, leaving them vulnerable when credit conditions deteriorate. Next, we make precise statements about inefficiency.

5.3 Implication III: Constrained Inefficiency

Section 3 we demonstrated that an inefficient organization of payments can lead to productive inefficiencies. In practice, governments may not be able to directly reorganize payments. However, they can influence aggregate expenditures via policy in-



(a) $\tilde{B}_t$ (b) B_t (c) $\Delta Y_t\%$

Figure 11: Resilience Comparison: Different values of δ.

Notes: The figure shows a numerical example of a credit crunch episode comparing two economies with different production discount factors ($\delta = 0.90$ vs. $\delta = 0.855$). Figures are computed via value-function iteration: $\beta = 0.90$, and $\tilde{B}_{ss} = 0.1 \cdot \bar{B}$. The economy has an initial debt of $B_0 = \tilde{B}_{ss}$ (outside the steady state). The crunch occurs in period t=6, when $\tilde{B}_t$ drops to $0.975 \cdot \tilde{B}_{ss}$. Then $\tilde{B}_t$ returns to steady state according to an AR(1) process with coefficient 0.9. The transition ends at t=41 when $\tilde{B}_t = \tilde{B}_{ss}$.

terventions.[52] This section studies a planner problem that respects the payments technology and financial constraints, but can determine the evolution of debt. The analysis allows us to isolate the expenditure externalities that differ from the organizational inefficiencies germane to the payments technology.

The Primal Planner Problem. We focus on constrained inefficiency during credit crunch episodes such as the one above. Recall that a credit crunch starts from an undisrupted steady state and ends in another undisrupted steady state. Since the economy is Pareto and productively efficient in an undisturbed steady state, Pareto inefficiencies can arise, if at all, only during a transition path—*although a distorted steady-state can also be constrained inefficient.* To study constrained inefficiencies within a transition, we consider a social planner with assigned Pareto weights that guarantee the same allocation in the terminal undisrupted steady state. This guarantees that the planner does not have redistributive objectives in the long run that would cause its chosen path of debt to deviate from that of the market outcome. This benchmark requires setting Pareto weights on entrepreneurs, θ, so that

$$\frac{S_{ss}}{C_{ss}} = \frac{1-(1-\beta)\,B_{ss}}{(1-\beta)\,B_{ss}} = \frac{\theta}{1-\theta}, \tag{14}$$

where B_{ss} is the terminal steady-state debt of the market equilibrium after the credit crunch.

We analyze a primal planner problem who directly chooses debt sequences $\{B_t\}_{t\geq 0}$ to maximize welfare, respecting the transactions technology and payment-chain frictions. Any allocation in this problem can be implemented with an appropriate sequence of debt and entrepreneur income taxes—see Appendix F.4.

Problem 4. (Primal Planner Problem):*Taking* $\left\{\tilde{B}_t\right\}$ *as given:*

$$\max_{\{B_t\}_{t\geq 0}} \sum_{t\geq 0} \beta^t \mathcal{P}\left(B_t, \tilde{B}_t\right)$$

where

$$\mathcal{P}\left(B, \tilde{B}\right) \equiv (1-\theta)\log\left((1-\beta)B\right) + \theta \log\left(S\left(B, \tilde{B}\right) + X\left(B, \tilde{B}\right)\right), \tag{15}$$

[52] Recall that relative to a random assignment, if it could, a planner would reorganize payments so that each chain is of equal length. Traffic regulation offers an analogy: governments cannot assign drivers to fast and slow lanes, but they can tax vehicle purchases to influence traffic.

and taking as given $S\left(B, \tilde{B}\right)$ and $X\left(B, \tilde{B}\right)$ as defined by Proposition 4.

For a given debt, by using the conditions in Proposition 4, the constrained planner respects the agents' expenditure decisions and input market-clearing conditions. Hence, the constrained planner respects the payment-chain technology (and production resource constraint) indirectly through the functions $S(B, \tilde{B})$ and $X(B, \tilde{B})$.

As the objective is time-separable[53] the optimal debt at t depends only on the current SBL $\tilde{B}_t$:

$$B_t = B^p(\tilde{B}_t) = \arg\max_{B \in [0, \bar{B}]} \mathcal{P}\left(B, \tilde{B}_t\right).$$

The solution is characterized in Appendix F. We summarize a key insight here. The solution reveals that during credit crunches, debt or its counterpart, interest rates, may be either inefficiently high or low. The reason is that the planner has two ways to induce more spot expenditures to increase TFP: it can redistribute wealth toward entrepreneurs (reducing their debt to free up spot borrowing) or toward creditors (increasing debt to raise their wealth) to stimulate their spot expenditures.

This ambivalent nature leads to four qualitatively different regimes depending on the tightness of the SBL, $\tilde{B}$. When $\tilde{B}$ is sufficiently ample, the economy operates efficiently with $\mathcal{A} = 1$. This is possible only if $\tilde{B}$ is sufficiently large so that at $B = B_{ss}$, spot expenditures $S(B, \tilde{B}) = 1 - (1 - \beta)B$ are feasible. As $\tilde{B}$ falls below that efficiency threshold, the planner faces a trade-off between productive efficiency and redistribution of resources toward entrepreneurs to increase their relative consumption in line with the Pareto weights. In moderate-constraint regions, the planner fully sacrifices redistribution to maintain the productive efficiency goal.[54] It does so by reducing debt, redistributing wealth one for one with $\tilde{B}$, guaranteeing that the entrepreneur only spends spots, but distorting its unconstrained consumption distribution, dictated by (14).

As the SBL falls further, the planner accepts some productive inefficiency, allowing the entrepreneur to make chained expenditures, but still reducing debt with the reduction in $\tilde{B}$, again redistributing from creditors toward entrepreneurs, balancing productive efficiency gains against worse redistribution. Most strikingly, when constraints become extreme, the planner reverses strategy entirely: rather than easing the

[53]The objective is time-separable because current debt determines creditor expenditures through log utility, with entrepreneur expenditures following residually.

[54]In typical planner problems, the objective is smooth. Hence, the planner typically balances objectives. Here, the planner forgoes redistribution of objectives, a corner solution that results from the discontinuity of marginal TFP losses from chained expenditures.

debt burden of constrained entrepreneurs, the planner redistributes wealth toward the *unconstrained agent*, the creditor. It essentially gives up on achieving productive efficiency by reducing debt and reverses course. This seemingly paradoxical result occurs because marginal increases in creditor wealth generate more spot expenditures which increases efficiency.

This ambivalent nature distinguishes payment-chain crises from macro-finance models, where redistributive and productive efficiency typically reinforce each other. Here, the planner may switch between complementary and conflicting policies as constraint severity varies because productive efficiency can be increased by redistributing wealth in either direction. Indeed, the planner's problem is generically not convex, thus, we cannot contrast the planner's first-order conditions with the aggregate Euler equation (13) to assess the inefficiency as typically done in such models.

Returning to Figure 10, illustrates how the planner's solution and the competitive equilibrium differ during a payment-chain crisis, demonstrating the inefficient debt dynamics. Notice how in the extreme phase, the planner implements regressive policies increasing debt, which cuts back the chained expenditures of entrepreneurs—and increasing spot expenditures of creditors. Notice also how this regressive policy induces significant TFP gains, while barely affecting the entrepreneur's consumption, making it almost a strictly Pareto-improving policy. In the hysteresis phase, the wealth is redistributed toward entrepreneurs who now have room to make spot expenditures. Whereas TFP remains depressed for a while, TFP recovers substantially. The planner sacrifices a slight reduction in the creditor's consumption for a substantial increase in TFP. In the deleveraging phase, both solutions converge, indicating the productive and Pareto efficiency of both solutions in the long run.

5.4 Implication IV: Fiscal Policy (Bocola Effect)

In this section, we study fiscal policy and show that government spending has strikingly different effects depending on when it is paid and on the state of the economy. The key is that when governments pay upfront (spot), they inject immediate liquidity into payment chains. When governments spend via chained expenditures, they extend payment chains and worsen delays. We call this the Bocola effect.[55]

To formalize this distinction, consider spot government expenditures G^s that require immediate payment, analogous to household spot purchases. The government

[55] Economist Luigi Bocola conjectured this result would hold in our setting.

can also make chained expenditures G^x contingent on receiving tax revenues, treating tax receipts as income flows. For simplicity, assume the government raises income taxes on entrepreneurs to finance expenditures under a balanced budget constraint. Government spending represents resource waste. The Government enters as another agent with a budget constraint similar to that of entrepreneurs (with B_t=0). Chained expenditures now include those of the government:

$$\mu_t = q_t X_t^w + G_t^x.$$

For the same level of government expenditures, spot government expenditures reduce the chained expenditure ratio by substituting for private chained purchases, while chained government expenditures increase μ_t.

Proposition 5. (Government Spending Multipliers): *Consider an infinitesimal increase in government expenditures evaluated at* $(G^x, G^s) = (0,0)$*. Consumption and output multipliers are:* ***Efficient Region*** *($B \leq B^*(\tilde{B})$): Both expenditure types reduce welfare:*

$$\frac{\partial S}{\partial G^x} = \frac{\partial S}{\partial G^s} = -1, \quad \frac{\partial \mathcal{Y}}{\partial G^s} = 0, \quad \frac{\partial \mathcal{Y}}{\partial G^x} = \mathcal{A}(\mu)(1+\epsilon_\mu^{\mathcal{A}}) - 1 < 0.$$

Crisis Region *($B > B^*(\tilde{B})$): Spot and chained expenditures have opposite effects:*

$$\frac{\partial X}{\partial G^x} = -1, \quad \frac{\partial X}{\partial G^s} = -(1+\epsilon_\mu^{\mathcal{A}}), \quad \frac{\partial \mathcal{Y}}{\partial G^x} = 0, \quad \frac{\partial \mathcal{Y}}{\partial G^s} = 1 - \mathcal{A}(\mu)(1+\epsilon_\mu^{\mathcal{A}}) > 0.$$

In words, the proposition states that in efficient regions, where no chained expenditures are made by the private sector, either form of government expenditures crowds out the entrepreneur's spot consumption. If the government expenditures are spot, they crowd out the entrepreneur's spot consumption. If the government spends through chained goods, it delays payments and reduces output, bearing all the losses directly.

During payment-chain crises in which the private sector makes chained expenditures, the effects differ. First, spot expenditures by the private sector do not change. Chained government expenditures remain harmful because they crowd out entrepreneur expenditures one-for-one. However, spot government expenditures create positive externalities by reducing average chain length. This shows a partial crowding in effect captured by $\epsilon_\mu^{\mathcal{A}}$, which is always negative. Moreover, when $\epsilon_\mu^{\mathcal{A}} < -1$, as in a deep crisis, it may be that spot government spending actually increases entrepreneur consumption despite higher taxes. The mechanism works through the speed-up of production, which more than offsets the cost of government resources.

This analysis yields a novel perspective on fiscal stimulus. The conventional view is that a government that satisfies budget balance, as in this exercise, can increase output by stimulating aggregate demand in the presence of price rigidities and no tax distortions. Here, we stress that with payment delays, it is important that the government pays upfront for goods. A government that buys goods from firms that need funds to accelerate payments may inadvertently reduce output efficiency. Effective stimulus requires paying upfront to break payment chains rather than extend them. This dimension differs fundamentally from standard aggregate demand arguments—or Ricardian considerations. Instead, fiscal policy works through a payments channel: it stimulates output only when it has positive externalities on the private sector's payments. Consistent with this novel mechanism, Barrot and Nanda (2020) finds evidence of a positive impact from faster government payments to small suppliers.

6 Conclusion

Financial crises have traditionally been understood through the lens of coordination failures and demand externalities. The contribution of this paper is to highlight payment delays as a contributing factor. We present empirical evidence that many firms rely on incoming payments to execute their transactions, and demonstrate, through an event study, how payment disruptions propagate through these payment chains.

In our theory, the core economic problem is that credit constraints lead to an inefficient distribution of funds, causing payment delays that disrupt production. Crises in this framework feature pecuniary externalities that arise endogenously from the payment structure. Although our policy recommendations resemble those in environments with price rigidity, we underscore the key importance of accelerating payments.

For tractability, we make several key simplifications. We assume all transactions are bilateral and of identical size, and that delays occur within periods rather than across periods. We abstract from risk and inflation. Developing extensions with richer transaction sequences and heterogeneous delays is essential for quantitative applications. Nonetheless, we believe the core lessons about payment-chain disruptions are robust to such generalizations.

References

Aguiar, Mark, Manuel Amador, Emmanuel Farhi, and Gita Gopinath, "Coordination and Crisis in Monetary Unions," *The Quarterly Journal of Economics*, 2015, *130* (4), 1727–1779. 4.1

Alvarez, Fernando and Gadi Barlevy, "Mandatory disclosure and financial contagion," *Journal of Economic Theory*, 6 2021, *194*, 105237. 1

Alvarez, Fernando E, David Argente, Francesco Lippi, Esteban Méndez, and Diana Van Patten, "Strategic Complementarities in a Dynamic Model of Technology Adoption: P2P Digital Payments," Working Paper 31280, National Bureau of Economic Research May 2023. 7

Argente, David, Esteban Méndez, and Diana Van Patten, "Cross-Border Product Adoption: Individual Imports, Migrant Networks, and Domestic Retailers," 2025. 7

Barrot, Jean-Noel, "Trade Credit and Industry Dynamics: Evidence from Trucking Firms," *The Journal of Finance*, 2016, *71* (5), 1975–2016. 3

Barrot, Jean-Noël and Ramana Nanda, "The Employment Effects of Faster Payment: Evidence from the Federal Quickpay Reform," *The Journal of Finance*, 2020, *75* (6), 3139–3173. 5.4

Biais, Bruno and Christian Gollier, "Trade Credit and Credit Rationing," *The Review of Financial Studies*, October 1997, *10* (4), 903–937. 3

Bigio, Saki and Jennifer La'O, "Financial Frictions in Production Networks," 2013. 2

Bocola, Luigi, "Trade credit and financial amplification," 2022. Unpublished Manuscript. 1, 38

Boissay, Frederic and Reint Gropp, "Trade Credit Defaults and Liquidity Provision by Firms," 2007. European Central Bank working paper series, 753. 3

Brunnermeier, Markus K. and Yuliy Sannikov, "The I-Theory of Money," 2017. 1

Buera, Francisco J. and Yongseok Shin, "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 2013, *121* (2), 221–272. 4.1

Cerdas, Ana María and Carlos Melegatti, *El sistema financiero costarricense en los últimos 25 años: sistema de pagos* number 2. In 'Monografía.', San José, Costa Rica: Academia de Centroamérica, 2014. 28 cm. 11

Cole, Harold L. and Timothy J. Kehoe, "Self-Fulfilling Debt Crises," *The Review of Economic Studies*, 2000, *67* (1), 91–116. _eprint: http://restud.oxfordjournals.org/content/67/1/91.full.pdf+html. 4.1

Costello, Anna Marie, "Credit Market Disruptions and Liquidity Spillover Effects in the Supply Chain," *Journal of Political Economy, forthcoming*, 2020. 3

Diamond, Peter A., "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 1982, *90*, 881–894. 1

Eggertsson, Gauti B. and Paul Krugman, "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach," *The Quarterly Journal of Economics*, July 2012, *127* (3), 1469–1513. _eprint: https://academic.oup.com/qje/article-pdf/127/3/1469/30457060/qjs023.pdf. 1

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson, "Financial Networks and Contagion," *American Economic Review*, 10 2014, *104*, 3115–53. 1

_ , _ , and Matthew V. Leduc, "Supply Network Formation and Fragility," *American Economic Review*, August 2022, *112* (8), 2701–47. 1

Freeman, Scott, "The Payments System, Liquidity, and Rediscounting," *The American Economic Review*, 1996, *86* (5), 1126–1138. Publisher: American Economic Association. 2

_ , "The Payments System, Liquidity, and Rediscounting," *The American Economic Review*, 1996, *86* (5), 1126–1138. 2

Green, Edward J., "Money and Debt in the Structure of Payments," *Federal Reserve Bank of Minneapolis Quaterly Review*, 1999, *23*, 13–29. 2

Guerrieri, Veronica and Guido Lorenzoni, "Liquidity and Trading Dynamics," *Econometrica*, 2009, *77* (6), 1751–1790. 2

_ **and** _ , "Credit Crises, Precautionary Savings, and the Liquidity Trap*," *The Quarterly Journal of Economics*, 2017, *132* (3), 1427–1467. 1

_ , _ , Ludwig Straub, and Ivan Werning, "Macroeconomic Implications of Covid-19: Can Negative Supply Shocks Cause Demand Shortages?," *American Economic Review*, 2022. 4

Hardy, Bryan, Ina Simonovska, and Felipe E. Saffie, "Economic Stabilizers in Emerging Markets:The Case for Trade Credit," 2022. 1

Jordà, Òscar, "Estimation and inference of impulse responses by local projections," *American economic review*, 2005, *95* (1), 161–182. 2, A.1

_ **and Alan M Taylor**, "Local projections," *Journal of Economic Literature*, 2025, *63* (1), 59–110. 2, A.1

Kiyotaki, Nobuhiro and John Moore, "Credit Chains," 1997. Publisher: London School of Economics. 1, 2

_ **and** _ , "Credit Cycles," *Journal of Political Economy*, 1997, *105* (2), 211–248. Publisher: The University of Chicago Press. 5

_ **and Randall Wright**, "On Money as a Medium of Exchange," *Journal of Political Economy*, 1989, *97* (4), 927–954. Publisher: The University of Chicago Press. 1, 2

Kopytov, Alexandr, Bineet Mishra, Kristoffer Nimark, and Mathieu Taschereau-Dumouchel, "Endogenous Production Networks under Supply Chain Uncertainty," 2022. 1

Lagos, Ricardo and Guillaume Rocheteau, "Liquidity in Asset Markets with Search Frictions," *Econometrica*, 2009, *77*, 403–426. 1

_ **and Randall Wright**, "Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy*, 2005, *113* (3), 463–484. 1, 2

_ **, Guillaume Rocheteau, and Pierre-Olivier Weill**, "Crises and Liquidity in OTC Markets," *Journal of Economic Theory*, 2011, *146*, 2169–2205. 1

La'O, Jennifer, "A Traffic Jam Theory of Recessions," 2015. 2

Li, Yiting, Guillaume Rocheteau, and Pierre-Olivier Weill, "Liquidity and the Threat of Fraudulent Assets," *Journal of Political Economy*, 2012, *120*, 815–846. 1

Lippi, Francesco, Stefania Ragni, and Nicholas Trachter, "Optimal monetary policy with heterogeneous money holdings," *Journal of Economic Theory*, 2015, *159*, 339 – 368. 1

Ljungqvist, Lars and Thomas J. Sargent, *Recursive Macroeconomic Theory*, 4th ed., Cambridge, MA: MIT Press, 2018. 2

Lucas, Robert E. and Nancy L. Stokey, "Money and Interest in a Cash-in-Advance Economy," *Econometrica*, 1987, *55* (3), 491–513. Publisher: The Econometric Society. 1, 3.2

Meza, Felipe and Erwan Quintin, "Factor Utilization and the Real Impact of Financial Crises," *The B.E. Journal of Macroeconomics*, 2007, *7* (1). 1

Méndez, Esteban and Diana Van Patten, "Voting on a Trade Agreement: Firm Networks and Attitudes Towards Openness," *The Review of Economic Studies*, 03 2025, p. rdaf005. 7

Nosal, Ed and Guillaume Rocheteau, *Money, Payment, and Liquidity*, MIT Press, 2011. 1

Oberfield, Ezra, "Productivity and misallocation during a crisis: Evidence from the Chilean crisis of 1982," *Review of Economic Dynamics*, 2013, *16* (1), 100–119. Special issue: Misallocation and Productivity. 1

_ , "A Theory of Input-Output Architecture," *Econometrica*, 3 2018, *86* (2), 559–589. 1

Pratap, Sangeeta and Carlos Urrutia, "Financial frictions and total factor productivity: Accounting for the real effects of financial crises," *Review of Economic Dynamics*, 2012, *15* (3), 336–358. 1

Rocheteau, Guillaume, "Payment and Liquidity under Adverse Selection," *Journal of Monetary Economics*, 2011, *58* (191-205). 1

— **, Tsz-Nga Wong, and Pierre-Olivier Weill**, "Working through the Distribution: Money in the Short and Long Run," May 2016. 1

Shepp, L. A. and S. P. Lloyd, "Ordered Cycle Lengths in a Random Permutation," *Transactions of the American Mathematical Society*, 1966, *121* (2), 340–357. 30

Shi, Shouyong, "A Divisible Search Model of Fiat Money," *Econometrica*, 1997, *65*, 75–102. 1

Taschereau-Dumouchel, Mathieu, "Cascades and Fluctuations in an Economy with an Endogenous Production Network," 2022. 1

Townsend, Robert M., "Models of Monetary Economies," 1980, pp. 265–303. 1

— , "Financial Structure and Economic Activity: Turnpike Theorems and Policy Implications," *Journal of Economic Theory*, 1983, *30* (2), 221–252. 2

von Schedvin, Erik Jacobsen Tor;, "Trade Credit and the Propagation of Corporate Failure: An Empirical Analysis," *Econometrica*, July 2015, *83* (4), 1315–1371. 3

Woodford, Michael, "Effective Demand Failures and the Limits of Monetary Stabilization Policy," *forthcoming American Economic Review*, 2022. 4

Appendix (Not intended for publication)

Table of Contents

A Appendix to the Empirical Analysis - Section 2

A.1 Heterogeneous Effects and Timing

Table A.1: Heterogeneous Effect by Size of Amounts Sent/Received

Dependent variable: Prob. firm j makes a payment at time t

	(1)
$received_{j,t-1} \times smaller\ sent_{jt}$	0.526
	(0.002)***
$received_{j,t-1}$	-0.109
	(0.001)***
Adjusted R^2	0.535
Observations	10,001,248
clusters	222,306
$t, j \times$ d FE	Yes

Notes: The independent variable $received_{j,t}$ is an indicator equal to one if firm j received a transfer from another firm within the last 24 hours. This variable is interacted with $smaller\ sent_{jt}$, which equals one if the amount received by firm j at time $t-1$ is less than the amount sent by firm j at time t, and zero otherwise. Robust standard errors, clustered by firm, are in parentheses. We include time and firm$\times$weekday fixed-effects. Data is daily and spans 2024.

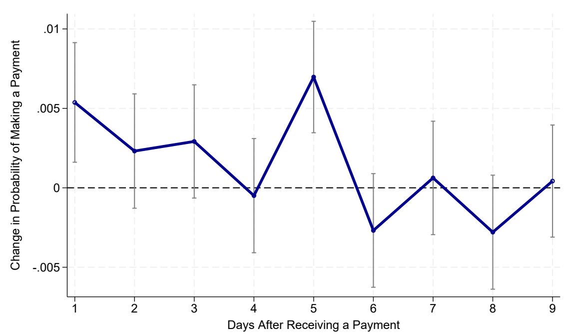


Figure A.1: Triggered Payments: Local Projection

Notes: The figure shows a local projection (Jordà, 2005; Jordà and Taylor, 2025), where the dependent and independent variables have 13 lags, as determined by a VAR-based optimal lag selection criteria. Observations are weighted by sales. Data is monthly and spans 2024.

A.2 Intensive Margin Regressions

We also study the intensive margin relationship between payments made and received. We follow equation (1), but now considering the (inverse hyperbolic sine of the) amounts sent and received by firm j. Results are presented in Table A.2 consistent with those of the extensive margins; firms become more likely to make payments once they receive money transfers. The elasticity across all firms in column (1) implies that a 1% increase in the amounts received by a firm on a given day increases the amounts sent the next day by approximately 0.04%.

Table A.2: Triggered Payments Across Firms—Intensive Margin

Dependent variable: Payments made by firm j at time t (asinh)

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln amount\, received_{j,t-1}$	0.036	0.017	0.045	0.044		0.036
(*asinh*)	(0.0005)***	(0.001)***	(0.001)***	(0.001)***		(0.0005)***
Small$_j$ $\times \ln amount\, received_{j,t-1}$		0.016				
		(0.001)***				
$\frac{Assets_j}{Sales_j} \times \ln amount\, received_{j,t-1}$			-0.019			
			(0.001)***			
$\frac{Liabilities_j}{Sales_j} \times \ln amount\, received_{j,t-1}$				-0.018		
				(0.001)***		
$\ln amount\, received_{j,t-5}$					0.012	0.012
(*asinh*)					(0.0004)***	(0.0004)***
Adjusted R^2	0.496	0.464	0.497	0.497	0.495	0.496
Observations	10,001,248	5,290,408	8,399,599	8,399,599	10,001,248	10,001,248
Clusters	222,306	74,566	174,896	174,896	222,306	222,306
$t, j \times$ d FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The independent variable in column (1) is the total amount (in colones) received by firm j from another firm within the last 24 hours, transformed using the inverse hyperbolic sine function. This variable is interacted with firm j's size, assets-to-sales and liabilities-to-sales in columns (2), (3), and (4) respectively. Columns (5) and (6) add an independent variable equal to the total amount (in colones) received by firm j from another firm 5 days before time t, transformed using the inverse hyperbolic sine function. Robust standard errors, clustered by firm, are in parentheses. We include time and firm$\times$weekday fixed-effects. Data is daily and spans 2024.

A.3 Placebo Exercises

Table A.3: Placebos—Receiving Payments Right After Making a Transfer

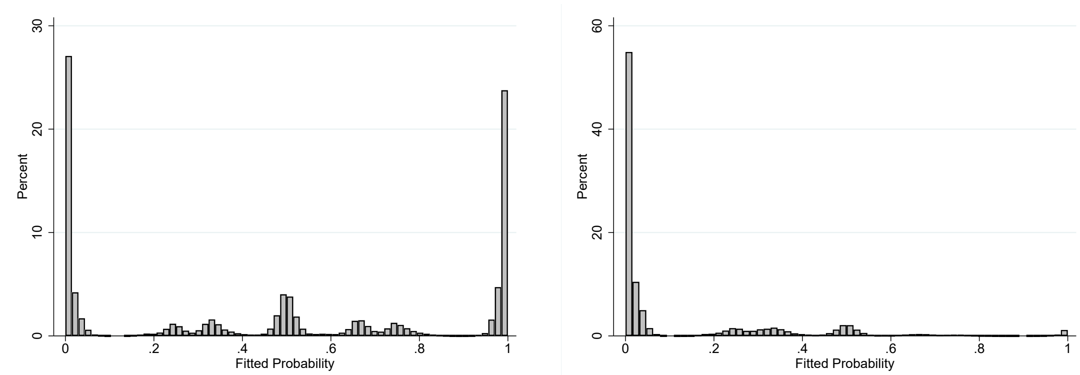
Dependent variable: Payment made by firm j at time t

	A. Extensive margin				B. Intensive margin (*asinh*)		
	$h = 1$	$h = 2$	$h = 3$		$h = 1$	$h = 2$	$h = 3$
	(1)	(2)	(3)		(4)	(5)	(6)
$received_{j,t+h}$	-0.001	-0.003	-0.001	$\ln amount$	-0.001	-0.002	0.000
	(0.0004)***	(0.0004)***	(0.0004)	$received_{j,t+h}$	(0.0004)	(0.0004)***	(0.0004)
Adjusted R^2	0.483	0.483	0.483		0.495	0.495	0.495
Observations	10M	10M	10M		10M	10M	10M
Clusters	222,306	222,306	222,306		222,306	222,306	222,306
$t, j \times$ d FE	Yes	Yes	Yes		Yes	Yes	Yes

Notes: In panel A, the independent variables are indicators equal to one if firm j received a payment in $t + h$. In panel B, the independent variables are the total amounts (in colones) received by firm j from another firm in $t + h$, transformed using the inverse hyperbolic sine function. Robust standard errors, clustered by firm, are in parentheses. We include time and firm$\times$weekday fixed-effects. Data is daily and spans 2024.

A.4 Results on Firm Classification and Chains

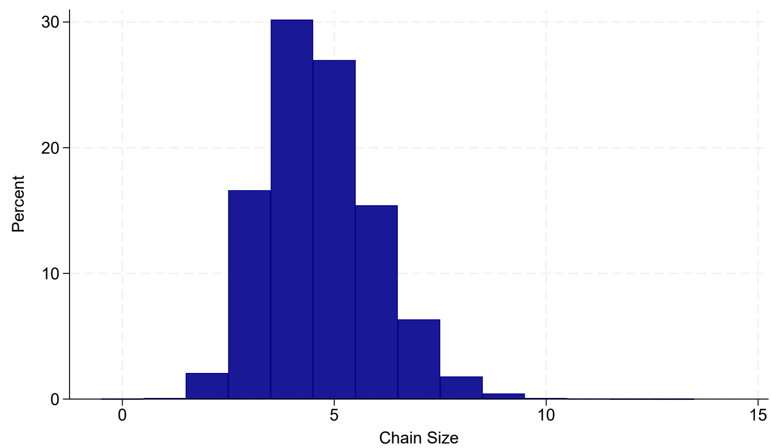
Figure A.2: Histogram of Fitted Probabilities of Triggered Payments by Firm



(a) Payments within 24 hours (b) Payments five days after

Notes: The figures show the distribution of monthly average triggered payments by firm based on an extensive margin estimation. Firms with a value of zero have a payment behavior that is orthogonal to receiving money transfers, while firms with a value of one have a perfect correlation between receiving a transfer and making a payment either the day after (panel a) or five days later (panel b). Data is monthly and spans 2024.

Figure A.3: Distribution of Chain Sizes



Notes: The figure shows the distribution of chain sizes based on a 5% random sample of spot firms in the economy used as initial nods. Data is collapsed for December 2024 to construct the chains.

Table A.4: Classification of Transactions by Month

Month	S→S	X→X	Mixed	Total	Month	S→S	X→X	Mixed	Total
1	1%	72%	28%	277,758	7	1%	72%	28%	291,866
2	1%	71%	28%	278,278	8	1%	72%	27%	276,308
3	1%	72%	28%	262,887	9	1%	72%	28%	280,320
4	1%	72%	28%	291,672	10	1%	71%	28%	290,711
5	1%	72%	28%	293,231	11	1%	71%	28%	282,262
6	1%	72%	28%	267,484	12	1%	72%	27%	287,585

Notes: The table shows the stability of the classification of transactions across months, depending on the buyer's and seller's type. Link types are divided into spot-to-spot (S→S), chained-to-chained (X→X), or mixed (S→X or X→S).

A.5 Chained Payments and Prices

We consider the following regression to explore the role of deferred payments on prices:

$$\ln Price_{vpjt} = \delta + \gamma\, Deferred_{vpjt} + \lambda_p + \lambda_j + \lambda_t + \varepsilon_{vpjt}$$

for transaction v, where $Deferred_{vpjt} = 1$ if sale v has a deferred payment and zero otherwise, and we include product, firm, and time fixed effects. We can also interact $Deferred_{vpjt}$ with an indicator, $Chained_{kt} = 1$, which depends on whether firm k—which made payment v to firm j—was classified as chained.

Column (1) of Table A.5 shows that prices for a given product are 21% higher in transactions with a deferred payment. Column (2) shows this effect is larger for firms classified as chained. Column (3) further breaks down transactions depending on the type of both the receiver (j) and sender (k) of the payment. Reassuringly, a higher price is charged only if the payment is made by a chained firm, regardless of the receiver's type.

Table A.5: Chained Payments and Prices

Dependent variable: Price of transaction v for good p sold by firm j to firm k at time t (log)

	(1)	(2)	(3)
$Deferred_{vpjt}$	0.189	0.111	0.181
	(0.006)***	(0.006)***	(0.006)***
Chained$_{kt}$ × $Deferred_{vpjt}$		0.091	
		(0.006)***	
Chained$_{jt}$&*Chained*$_{kt}$ × $Deferred_{vpjt}$			0.065
			(0.002)***
Spot$_{jt}$&*Chained*$_{kt}$ × $Deferred_{vpjt}$			0.058
			(0.003)***
Spot$_{jt}$&*Spot*$_{kt}$ × $Deferredt_{vpjt}$			0.006
			(0.004)
Adjusted R^2	0.682	0.682	0.682
Observations	45,801,998	45,801,998	45,801,998
Clusters	175,829	175,829	175,829
p, j, t FE	Yes	Yes	Yes

Notes: The independent variable in column (1) is an indicator equal to 1 if transaction v;s payment was deferred, as opposed to having an immediate payment. This variable is interacted with a dummy that depends on firm k's classification as chained or spot in column (2). Column (3) further decomposes transactions depending on whether firm j and k were chained or spot at t; the omitted category has firm j as chained and firm k as spot. Robust standard errors, clustered by firm, are in parentheses. We include product, firm, and time fixed-effects. Data is by sale and spans 2024.

A.6 Payments and Timing of Delivery

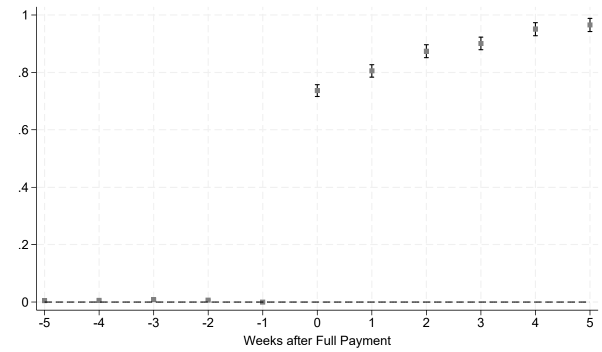
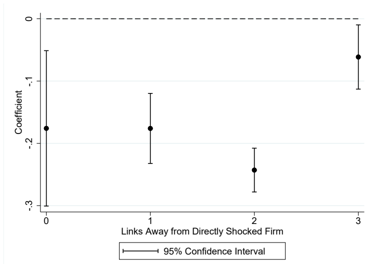


Figure A.4: Timing of Product Delivery

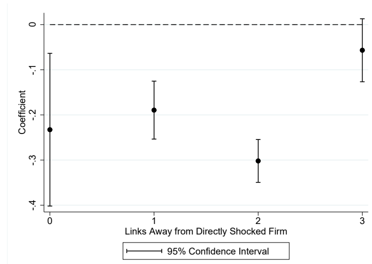
Notes: The figure shows an event study for orders with deferred payment in which product sales become positive after the *full* payment is due. Time zero corresponds with the date in which the full payment is due and the coefficient represents the sales of the product at the store.

A.7 Persistence of Desyfin Effect Along the Chain

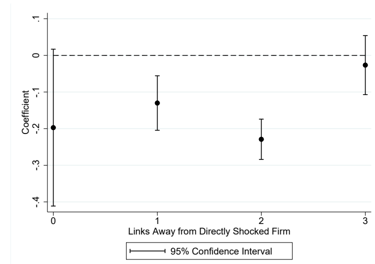
Table A.6: Desyfin Effect Along the Chain: Firms Upstream From Directly Hit Firm



(a) $\ln Purchases_{jt}$



(b) $\ln Sales_{jt}$



(c) $\ln Value\ Added_{jt}$

Notes: The figure shows how the impact of being upstream from a Desyfin firm evolves as increase the number of links of separation along the chain. The dependent variables in panels (a), (b), and (c) are the (log) of purchases, sales, and value added, respectively, of firm j in month t. The plotted coefficients correspond with the triple interaction of $Links(n)_{jt} \times Switch2chained_{it} \times Desyfin_i \times Shock_t$, and captures the impact of being n links away from a Desyfin firm i ($Links(n)_{jt}$) *and* firm i switching to being classified as chained *after* the freeze. Robust standard errors are clustered by firm. We include province×industry (ISIC-4) and time fixed-effects. Data is monthly and spans 2024.

B Appendix to Payment-Chain Block (Section 3)

B.1 Microfoundations for formula for y_n

B.1.1 Value Added - Delayed Production

The goal of this section is to provide a setting where value added for an order in position n is given by $y_n = \delta^n$ and the transaction time is $\tau_n = 1 - \delta^n$. This formulation ensures that all transactions occur within the unit time interval $[0, 1]$, with $\tau_n \to 1$ as $n \to \infty$. This within-period structure allows us to embed payment chains into an infinite-horizon discrete-time framework, as we do in Section 5. Specifically, the entire payment-chain network—from the first spot order at $\tau_0 = 0$ to the limiting transaction as $n \to \infty$—forms and settles within each integer period. This approach preserves analytical tractability while capturing the essential dynamics of sequential payments and production delays.

Time and Production. Value added is generated by translating inputs into final goods. To produce final goods, input endowments must be transferred to the agent placing the order for the good. Recall that there are N orders in the production network, indexed by their position $k = 1, 2, \ldots, N$ in the bilateral production network, induced by an assignment rule $\mathcal{P}$, as in the main text. Within this production network, payment chains emerge. Further recall that we index orders by their position $n = 0, 1, 2, \ldots$ within a payment chain. An order at production position k corresponds to some position n in its payment chain.

Value added. As in the main text, orders are for customized intermediate inputs: An order at position k in the bilateral production network is supplied by an order in position $k + 1$ and so on. The customized nature of these goods means they have no resale value. Moreover, recall that the production network and the income-expenditure identity induce a network of payment chains.

An order in position n in a payment chain is fulfilled at a time $\tau_n \in [0, 1]$. This means that the intermediate input is available to the agent placing the order at τ_n. The technology to produce final goods (gross output) is given by:

$$y_n = \int_{\tau_n}^{1} \min\{x_t, h_t\} dt. \quad (16)$$

Here, h_t stands for the labor input per unit of time that can be used by the agent placing the order to produce final goods. We assume that the agent placing the order has an endowment of labor $h_t \leq 1$ to process each order.[56] In turn, x_t represents the use of the intermediate input per unit of time. The total input endowment is 1 unit, and input use must satisfy:

$$\int_{\tau_n}^{1} x_t dt \leq 1. \tag{17}$$

Because once payments for orders have been executed and there's no resale use for the inputs, the only decision concerns the use of h_t and x_t. To maximize value added y_n, requires setting $h_t = x_t = 1$. Thus, output is given by $y_n = \int_{\tau_n}^{1} \min\{x_t, h_t\} dt = 1 - \tau_n$. Suppose that transaction times indeed occur at times $\tau_n = 1 - \delta^n$. Thus, in that case, $y_n = \delta^n$.

These calculations imply that for an order in position n, the share $\tau_n = 1 - \delta^n$ of its input is wasted.[57] Next, we derive an explanation of why the transaction times are indeed instances $\tau_n = 1 - \delta^n$.

Delays arising from fraud prevention. We consider a scenario where delays in transferring funds and, therefore, delays in transferring inputs are the result of a payment system designed to deter fraud.

To that end, we consider the possibility of fraudulent orders. On the one hand, we have *buyer-side fraud*: a population of fraudster agents can place fraudulent chain orders. These fraudulent chained orders falsely claim to have incoming payments, but in reality, they have nothing to sell and know that no incoming funds will arrive. The presence of these agents implies that to prevent a flooding of buyer-side fraud, any transfer of inputs must require *proof of funds*; otherwise, trade would break down.

On the other hand, we consider the possibility of *seller-side fraud*: a population of fraudster agents that can provide fraudulent inputs. These inputs generate no value added. However, by selling these inputs, fraudster sellers can obtain funds to buy inputs while harming their buyers. Trade is possible only if goods can be inspected, which takes time.

To enable trade despite these fraudulent orders and suppliers, transactions are conducted through a payment system that enables an escrow-inspection mechanism:

[56] We could alternatively model a competitive market for labor with elastic supply, which would not change the main results leading to the same outcome.

[57] In the extensions section, we discuss variations so that the unused inventory is not waste.

1. Once an order at position n receives funds at τ_n, the buyer can deposit them into an escrow mechanism. The funds can be proven to exist, but will not be released until the buyer sends out a payment instruction.

2. Once the funds are in escrow, the seller (at position $k+1$ in the production network) verifies that the buyer is not fraudulent. The seller delivers the inventory immediately. As a result, the buyer has the inventory at time τ_n. The buyer begins production immediately. The presence of verifiable funds deters buyer-side fraud.

3. The buyer requires inspection to deter seller-side fraud. To do so, the buyer must produce a fraction $(1-\delta)$ of output as a quality verification sample. This inspection requires to first observe the share $(1-\delta)$ of the output that will be produced.

4. The escrowed funds are released to the seller at position $n+1$ in the payment chain. If the seller proves the inputs are fraudulent, the funds are returned to the buyer. Whenever the full input units arrive in some $\tau > 0$ not all the input can be used. In particular, the buyer will use only share $x = 1 - \tau_n$ of the input. Thus, the inspection requires $(1-\delta)(1-\tau_n)$ inspection time given that the output is just $(1-\tau_n)$.

Notice that the parameter $\delta \in (0,1)$ represents inspection efficiency or verification technology. Furthermore, observe that if buyers are non-fraudulent, they execute payment. We are implicitly assuming commitment on the side of honest buyers: they pay the seller even when having the inventory and despite how late they may be in the chain.

Derivation of Timing of Payments. Consider position n in the chain: the buyer receives inventory at time τ_n. Since value added is: $y_n = 1 - \tau_n$ of which $(1-\delta)(1-\tau_n)$ is the inspection time, payment is released to position $n+1$ at: $\tau_n + (1-\delta)y_n$. Thus, we obtain the recursion:

$$\tau_{n+1} = \tau_n + (1-\delta)y_n = \tau_n + (1-\delta)(1-\tau_n) = (1-\delta) + \delta\tau_n \tag{18}$$

The initial condition is $\tau_0 = 0$ (spot orders begin immediately with external funding). Thus, we obtain the recursion:

$$\begin{aligned} \tau_1 &= 1 - \delta \\ \tau_2 &= (1 - \delta) + \delta(1 - \delta) = 1 - \delta^2 \\ \tau_3 &= (1 - \delta) + \delta(1 - \delta^2) = 1 - \delta^3. \end{aligned}$$

Hence, we establish the desired result that:

$$y_n = 1 - \tau_n = \delta^n. \tag{19}$$

Examples of escrow-inspection mechanisms in practice. Our setting is admittedly simplistic. However, the payment delays that occur through escrow-inspection mechanisms we model—involving proof of funds, escrow, and quality verification—are ubiquitous in real-world payment systems.

The clearest example arises in international trade, where customs agencies serve precisely this dual function. They hold shipments in bonded warehouses until payment confirmation is verified, ensuring proof of funds. Simultaneously, they inspect containers before releasing payment, verifying quality and compliance. This process creates exactly the delay structure in our model.

Traditional payment technologies, such as checks, embody a similar logic. When a buyer pays by check, processing delays provide a window for inspection before funds transfer. If goods prove defective, the buyer can stop payment; if the seller lacks sufficient backing, the check bounces, revealing fraud. The check-clearing delay thus naturally accommodates inspection before irrevocable settlement.

These mechanisms are equally central to digital commerce. Online marketplaces such as eBay and MercadoLibre operate buyer and seller protection programs that hold payments in escrow until buyers confirm receipt and quality. In business-to-business transactions, vendors commonly require bank confirmation letters proving funds exist before shipping goods, while buyers release final payment only after inspection and acceptance. Blockchain technologies and digital wallets extend these principles further: wallet balances are publicly verifiable, and customers can demonstrate ownership and proof of funds by initiating small test transactions, providing transparent verification to all parties in the chain.

These institutional arrangements reflect the fundamental force: delays in payment execution are the unavoidable cost of enabling transactions when two-sided fraud is possible. While we do not study optimal payment system design in this paper—an interesting topic in its own right—our framework captures a tension between transaction speed and fraud prevention.

B.1.2 Net Present Value with Delayed Transactions

We now turn to another variation. We now consider the payment times to be $\tau \in 0, 1, ..., \infty$. Instead, in the spirit of Townsend's original formulation, there is full production, but production is valued at δ^n from the perspective of the moment when chains are formed. In this case, the good is perfectly storable, so δ^n represents the discounted value of the stored unit.

B.2 Assignment Functions and Outcomes (Proof of Proposition 1)

Proposition 1 presents several results for random and efficient assignments. We start with the formulas for random assignments

B.2.1 Random Assignments.

Chain-Length Distribution. We first show that the distribution of payment-chain lengths is geometric.

Lemma 2. *For the random assignment, the chain-length distribution is geometric.*

Proof. Recall that the order assignment is random and nodes are either spot orders ($\mathcal{N}^s$) or chained ($\mathcal{N}^x$). As $N \to \infty$, each node has an outgoing link with probability μ (the order buys from a product committed to another chained order) or no incoming link with probability $1 - \mu$, the link is with a spot order.

By construction, every chained order must be funded through exactly one production link, creating a directed path that ultimately terminates at a spot order. This structure ensures that every node belongs to exactly one payment chain, and each chain originates from a unique spot order. Therefore, we have exactly N^s payment chains, which we index by $i \in \mathcal{N}^s$.

Because there is a one-to-one map from spot orders to chains, we use an index for all possible chains by defining chain i as the payment chain that starts with spot order i.

For chain i, corresponding to an initial spot order i, we define its length as the number of chained orders it contains (equivalently, the number of edges in the directed path). We now derive the distribution of chain lengths.

Consider chain i: It is of **length 0** if spot order i orders from a node associated with a spot order, which occurs with probability $1-\mu$. It is of **length 1** the first order in the chain orders from a chained order, which happens with probability μ, whereas that order is not linked, which occurs with probability $1-\mu$. Thus, the overall probability is $\mu(1-\mu)$. Proceeding by induction, for a chain of **length** n, extending this reasoning, there must be n consecutive chained orders, followed by a termination (no further links). By independence, this occurs with probability $\mu^n(1-\mu)$. Therefore, if L_i denotes the length of chain i, we have:

$$\mathbb{P}(L_i = n) = \mu^n(1-\mu) \quad \text{for } n \in \{0, 1, 2, \ldots\},$$

i.e., the geometric distribution with parameter $1-\mu$ (using the convention where support begins at 0). □

Output and TFP. Next, we derive the distribution of positions for the geometric distribution.

Lemma 3. *When the chain length distribution is geometric, the distribution of positions is also geometric.*

Proof. For the geometric distribution:

$$G\left(n-1\right) = \left(1-\mu\right)\sum_{m=0}^{n-1}\mu^n = \left(1-\mu^n\right), \quad n \geq 1.$$

Therefore, $1 - G\left(n-1\right) = \mu^n$, and, thus,

$$\sum_n \left(1 - G\left(n-1\right)\right) = \sum_{n=0}^{\infty}\mu^n = \frac{1}{1-\mu}.$$

Thus, we have that:

$$h\left(n\right) = \frac{\mu^n}{1/\left(1-\mu\right)} = \left(1-\mu\right)\mu^n.$$

Thus, the distribution of positions follows a geometric distribution with the same parameter $1-\mu$, completing the proof. □

Next, we derive the formula for TFP, $\mathcal{A}$, and output $\mathcal{Y}$ using two approaches. We first do so by employing the distribution of positions.

Lemma 4. *For the random assignment we have that* $\mathcal{Y} = \frac{1-\mu}{1-\delta\mu}$ *and* $\mathcal{A} = \delta\frac{1-\mu}{1-\delta\mu}$.

Calculations using the Position Distribution h.

Proof. Given that the distribution of positions $h(n) = (1-\mu)\mu^n$ the average quantity of final goods,

$$\mathcal{Y} \equiv \mathbb{E}[y_n] = \mathbb{E}[\delta^n] = \sum_{n=0}^{\infty}(1-\mu)\mu^n\delta^n = \frac{1-\mu}{1-\delta\mu}.$$

Then, it follows that $\mathcal{Y} = (1-\mu) + \mu\mathcal{A} \rightarrow \mathcal{A} = \delta\frac{1-\mu}{1-\delta\mu}$. □

Calculation using the Length Distribution.

Alternative Proof via Weighted Averaging. We demonstrate a general approach for deriving position distributions from chain length distributions, applicable even when direct computation is complex. We illustrate with the geometric case where $\mathbb{P}(L = n) = (1-\mu)\mu^n$.

Step 1: Expected production per chain. For a chain of length $n \geq 1$, the cumulative production discount is:

$$E[y_n|\text{length} = n] = \sum_{i=1}^{n}\delta^i = \frac{\delta(1-\delta^n)}{1-\delta}$$

where $\delta \in (0, 1)$ represents the production efficiency at each stage.

Step 2: Average chain length. The expected number of chained orders per chain (excluding spot orders) is:

$$\bar{n}_x = \sum_{n=1}^{\infty} n \cdot \mathbb{P}(L = n) = \sum_{n=1}^{\infty} n(1-\mu)\mu^n \tag{20}$$

$$= (1-\mu)\mu\sum_{n=1}^{\infty} n\mu^{n-1} \tag{21}$$

$$= (1-\mu)\mu \cdot \frac{d}{d\mu}\left(\frac{1}{1-\mu}\right) \tag{22}$$

$$= (1-\mu)\mu \cdot \frac{1}{(1-\mu)^2} = \frac{\mu}{1-\mu} \tag{23}$$

Step 3: Position distribution via weighted averaging. To find the distribution of positions, we weight each chain length by its contribution to the total mass of chained orders:

$$\text{Weight for length } n = \frac{n \cdot \mathbb{P}(L = n)}{\bar{n}_x}$$

The probability that a randomly selected chained order is at position k equals:

$$\mathbb{P}(X = k | X > 0) = \sum_{n=k}^{\infty} \frac{(1-\mu)\mu^n \cdot 1}{\mu/(1-\mu)} = (1-\mu)^2 \sum_{n=k}^{\infty} \mu^{n-1} = (1-\mu)\mu^{k-1}$$

Accounting for spot orders (position 0), we have:

$$\mathbb{P}(X = k) = \begin{cases} 1-\mu & \text{if } k = 0 \\ \mu \cdot (1-\mu)\mu^{k-1} = (1-\mu)\mu^k & \text{if } k \geq 1 \end{cases}$$

This confirms that positions follow a geometric distribution with parameter $1 - \mu$.

□

Part 2. Limits. We first consider the limit as $\mu \to 0$:

$$\lim_{\mu \to 0} \mathcal{A}(\mu; \delta) = \delta \lim_{\mu \to 0} \left(\frac{1-\mu}{1-\delta\mu} \right) = \delta.$$

For output:

$$\lim_{\mu \to 0} \mathcal{Y}(\mu; \delta) = \lim_{\mu \to 0} (1-\mu) + \mu\mathcal{A}(\mu; \delta) = 1.$$

Next, we consider the limit as $\mu \to 1$:

$$\lim_{\mu \to 1} \mathcal{A}(\mu; \delta) = \delta \lim_{\mu \to 1} \left(\frac{1-\mu}{1-\delta\mu} \right) = 0.$$

For output:

$$\lim_{\mu \to 1} \mathcal{Y}(\mu) = \lim_{\mu \to 1} (1-\mu) + \lim_{\mu \to 1} \mu \lim_{\mu \to 1} \mathcal{A}(\mu) = 0.$$

Next, we consider the limit as $\delta \to 0$:

$$\lim_{\delta \to 0} \mathcal{A}(\mu; \delta) = \lim_{\delta \to 0} \frac{\delta - \delta\mu}{1 - \mu\delta} = 0.$$

For output,

$$\lim_{\delta\to 0}\mathcal{Y}(\mu) = (1-\mu) + \mu\lim_{\delta\to 0}\mathcal{A}(\mu;\delta) = (1-\mu).$$

Finally, we consider the limit as $\delta \to 1$:

$$\lim_{\delta\to 1}\mathcal{A}(\mu;\delta) = \lim_{\delta\to 1}\delta\cdot\lim_{\delta\to 1}\frac{1-\mu}{1-\delta\mu} = 1.$$

For output,

$$\lim_{\delta\to 1}\mathcal{Y}(\mu) = (1-\mu) + \mu\lim_{\delta\to 1}\mathcal{A}(\mu;\delta) = 1.$$

Part 3. Monotonicity and concavity. We establish the following result:

Lemma 5. *$\mathcal{Y}(\mu), \mathcal{A}(\mu)$ are strictly decreasing and concave in μ.*

Proof. Next, we investigate the monotonicity of $\mathcal{A}$ and $\mathcal{Y}$. We do so by signing the derivatives.

$$\mathcal{Y}(\mu) = (1-\mu)\left(1+\frac{\delta\mu}{1-\delta\mu}\right) = \frac{1-\mu}{1-\delta\mu}.$$

Therefore, we obtain,

$$\begin{aligned}\mathcal{Y}_{\mu} &= \mathcal{Y}(\mu)\left(\frac{\delta}{1-\delta\mu}-\frac{1}{1-\mu}\right)\\ &= -\frac{1-\delta}{(1-\delta\mu)^2} < 0.\end{aligned}$$

We also have that, $\mathcal{Y}_{\mu} = -(1-\mathcal{A}(\mu)) + \mu\mathcal{A}_{\mu}$. Then,

$$1-\mathcal{A}(\mu) = \frac{1-\delta\mu-\delta+\delta\mu}{1-\delta\mu} = \frac{1-\delta}{1-\delta\mu}.$$

Therefore:

$$\mu\mathcal{A}_{\mu} = \left(\frac{1-\delta}{1-\delta\mu}-\frac{1-\delta}{(1-\delta\mu)^2}\right) = -\frac{1-\delta}{1-\delta\mu}\left(\frac{\delta\mu}{1-\delta\mu}\right) < 0.$$

Thus:

$$\mathcal{A}_{\mu} = -\frac{1-\delta}{1-\delta\mu}\left(\frac{\delta}{1-\delta\mu}\right). \tag{24}$$

Next, we perform the convexity analysis by computing second derivatives. Using (25) we have that:

$$\mathcal{Y}_{\mu\mu} = -2\delta\frac{1-\delta}{(1-\delta\mu)^3} < 0.$$

Thus, $\mathcal{Y}(\mu)$ is strictly concave.

From $\mathcal{Y}_{\mu\mu} = 2\mathcal{A}_\mu + \mathcal{A}_{\mu\mu}$, we obtain

$$\begin{aligned}\mathcal{A}_{\mu\mu} &= \mathcal{Y}_{\mu\mu} - 2\mathcal{A}_\mu \\ &= -2\delta\frac{1-\delta}{(1-\delta\mu)^3} + 2\frac{1-\delta}{1-\delta\mu}\left(\frac{\delta}{1-\delta\mu}\right) \\ &= 2\delta\frac{1-\delta}{(1-\delta\mu)^2}\left(\delta - \frac{1}{1-\delta\mu}\right) < 0.\end{aligned}$$

Thus, we establish that both functions are strictly monotonically decreasing and concave in μ.

□

We also have monotonicity in δ, but a different convexity property.

Lemma 6. *$\mathcal{Y}(\mu), \mathcal{A}(\mu)$ are strictly increasing in δ. In turn, $\mathcal{Y}(\mu)$ is strictly concave in δ whereas $\mathcal{Y}(\mu)$ is strictly convex in δ.*

Proof. Note that:

$$\mathcal{Y}_\delta = \mathcal{Y}(\mu)\left(\frac{\mu}{1-\delta\mu}\right) > 0.$$

Moreover, $\mu\mathcal{A}_\delta = \mathcal{Y}_\delta \to \mathcal{A}_\delta = \frac{\mathcal{Y}(\mu)}{1-\delta\mu} > 0.$

As for the convexity analysis:

$$\mathcal{Y}_{\delta\delta} = \mathcal{Y}_\delta\left(\frac{\mu}{1-\delta\mu}\right) + \mathcal{Y}(\mu)\frac{\mu^2}{(1-\delta\mu)^2} = 2\mathcal{Y}(\mu)\frac{\mu^2}{(1-\delta\mu)^2} > 0.$$

From here, we have that, $\mathcal{A}_{\delta\delta} = \mathcal{Y}_\delta\left(\frac{1}{1-\delta\mu}\right) + \mathcal{Y}(\mu)\left(\frac{\mu}{1-\delta\mu}\right) > 0$. This establishes the strict monotonicity and convexity properties. □

B.2.2 Efficient Assignment

We now prove the formulas for the chain-size distribution for the efficient assignment.

Lemma 7. *Let, $n = \left\lfloor \frac{\mu}{1-\mu} \right\rfloor$ and $p = n + 1 - \frac{\mu}{1-\mu}$. Then, under the efficient assignment* $g(n) = p, \quad g(n+1) = 1-p$.

Proof. Consider the problem of allocating $N^x = \mu N$ chained orders across $N^s = (1 - \mu)N$ chains to maximize total output.

Step 1: Marginal productivity and reallocation.

Consider any assignment with two chains of lengths n and n' where $n' > n + 1$. We can strictly increase output by moving the order from the final position in the chain of length n' to adding a link to the chain of length n. The final goods obtain by that order is:

$$\delta^{n+1} - \delta^{n'} = \delta^{n+1}(1 - \delta^{n'-n-1}) > 0.$$

The final goods of no other order is not affected.

Step 2: Characterization of efficient allocation. By Step 1, in any efficient allocation, chain lengths can differ by at most 1. Therefore, the distribution must be:

$$g(k) = \begin{cases} p & \text{if } k = n \\ 1 - p & \text{if } k = n + 1 \\ 0 & \text{otherwise} \end{cases}$$

for some integer $n \geq 0$ and $p \in [0, 1]$. Otherwise, it's possible to move a mass of orders and increase the final output.

Step 3: Determining n and p. The total number of chained orders must equal N^x:

$$N^s[pn + (1 - p)(n + 1)] = N^x.$$

Dividing both sides by N^s:

$$pn + (1 - p)(n + 1) = \frac{N^x}{N^s} = \frac{\mu N}{(1 - \mu)N} = \frac{\mu}{1 - \mu}.$$

Simplifying:

$$n + (1 - p) = \frac{\mu}{1 - \mu}.$$

Therefore:

$$p = n + 1 - \frac{\mu}{1 - \mu}.$$

Step 4: Integer constraints. For $p \in [0, 1]$, we need:

$$0 \leq n + 1 - \frac{\mu}{1-\mu} \leq 1. \tag{25}$$

This implies:

$$n \leq \frac{\mu}{1-\mu} < n + 1.$$

Hence, $n = \left\lfloor \frac{\mu}{1-\mu} \right\rfloor$, $n + 1 = \left\lceil \frac{\mu}{1-\mu} \right\rceil$, and $p = n + 1 - \frac{\mu}{1-\mu}$.
Step 5: Final distribution. The efficient allocation has:

- A fraction $p = \left\lceil \frac{\mu}{1-\mu} \right\rceil - \frac{\mu}{1-\mu}$ of chains with length $n = \left\lfloor \frac{\mu}{1-\mu} \right\rfloor$
- A fraction $1 - p = \frac{\mu}{1-\mu} - \left\lfloor \frac{\mu}{1-\mu} \right\rfloor$ of chains with length $n + 1$

Note that for $\mu < \frac{1}{2}$, the mass of chains of length 1 is $1 - p = \frac{\mu}{1-\mu} = \frac{N^x}{N^s}$, precisely the ratio of chained orders to spot orders.

This completes the characterization of the efficient chain-length distribution. □

The distribution of positions $h(n)$ is thus given by:

$$H(k) = \begin{cases} \frac{1}{n+(1-p)} & \text{if } k \leq n \\ \frac{(1-p)}{n+(1-p)} & \text{if } k = n+1 \\ 0 & \text{otherwise} \end{cases}$$

where $n = \left\lfloor \frac{\mu}{1-\mu} \right\rfloor$. Next, we derive the formula for $\mathcal{A}$ under the efficient assignment.

Lemma 8. *Under the efficient assignment:*

$$\mathcal{A}(\mu; \delta) = \delta \left(p \frac{1-\delta^n}{1-\delta} + (1-p) \frac{1-\delta^{n+1}}{1-\delta} \right) \frac{1-\mu}{\mu}.$$

Furthermore:

$$\mathcal{Y}(\mu; \delta) = (1-\mu) \left(1 + \delta \left(p \frac{1-\delta^n}{1-\delta} + (1-p) \frac{1-\delta^{n+1}}{1-\delta} \right) \right).$$

Proof. Note that in a chain of length n, the average output among chained orders: $\frac{1}{n} \sum_{i=1}^{n} \delta^i = \frac{\delta}{n} \frac{1-\delta^n}{1-\delta}$. The mass of chained orders is: $pn + (1-p)(n+1) = \frac{\mu}{1-\mu}$. From here, we compute output per chained order:

$$\mathcal{A}(\mu) = \delta \left(p \frac{1-\delta^n}{1-\delta} + (1-p) \frac{1-\delta^{n+1}}{1-\delta} \right) / \left(\frac{\mu}{1-\mu} \right), \tag{26}$$

yielding the formula. □

Monotonicity and concavity. Next we prove the established monotonicity and continuity.

Lemma 9. *$\mathcal{A}(\mu)$ and $\mathcal{Y}(\mu)$ are continuous and monotone decreasing under the efficient allocation.*

Proof. Recall that under efficient allocation, average output per chained order is:

$$\mathcal{A}(\mu) = \frac{\delta}{1-\delta} \left(p(1-\delta^n) + (1-p)(1-\delta^{n+1}) \right) \frac{1-\mu}{\mu}$$

where $n = \left\lfloor \frac{\mu}{1-\mu} \right\rfloor$ and $p = n + 1 - \frac{\mu}{1-\mu}$.

Step 1: Continuity at integer values. When $\frac{\mu}{1-\mu} = k$ for integer k, all chains have length k, giving:

$$\mathcal{A}(\mu) = \frac{\delta(1-\delta^k)}{1-\delta} \cdot \frac{1-\mu}{\mu}$$

As μ approaches this integer point from below, $n = k - 1$ and $p \to 0$, yielding the same limit. Thus $\mathcal{A}(\mu)$ is continuous everywhere.

Step 2: Derivative between integer points. For $\frac{\mu}{1-\mu} \notin \mathbb{Z}$, the floor function is locally constant, so:

$$\frac{\partial \mathcal{A}}{\partial \mu} = \frac{\delta}{1-\delta} \cdot \frac{\partial}{\partial \mu} \left(\frac{1-\mu}{\mu} \right) \cdot [(1-\delta^n) - n(1-\delta)\delta^n]$$

Since $\frac{\partial}{\partial \mu} \left(\frac{1-\mu}{\mu} \right) = -\frac{1}{\mu^2} < 0$, we need to show:

$$(1-\delta^n) - n(1-\delta)\delta^n > 0$$

Step 3: Verifying the inequality. Rearranging:

$$(1-\delta^n) - n(1-\delta)\delta^n = 1 - (1+n)\delta^n + n\delta^{n+1} > 0.$$

We need to prove: $1 > (1+n)\delta^n - n\delta^{n+1}$. Define $f(\delta) = 1 + n\delta^{n+1} - (1+n)\delta^n$. We have: $f(0) = 1 > 0$, $f(1) = 0$, $f'(\delta) = n(n+1)\delta^n - n(1+n)\delta^{n-1} = n(1+n)\delta^{n-1}(\delta - 1) < 0$. Since

f is strictly decreasing from positive to zero on $[0,1]$, we have $f(\delta) > 0$ for all $\delta \in (0,1)$. This verifies condition for local monotonicity.

Step 4: Global monotonicity. Although the derivative has discontinuities at integer values of $\frac{\mu}{1-\mu}$, the function is continuous and decreasing between these points. At integer points, the left and right limits coincide, ensuring global monotonicity.

Therefore, $\mathcal{A}(\mu)$ is strictly decreasing in μ for all $\mu \in (0,1)$.

Step 4: Monotonicity in δ. This follows immediately since δ^a is increasing in δ for any integer a.

Step 5: Concavity in μ. Since for any non integer value of $\frac{\mu}{1-\mu}$, the floor and ceiling functions and, thus, n, it suffices to show concavity in p. This requires to show convexity in $\frac{\mu}{1-\mu}$. The second derivative of this ratio is $\frac{1}{1-\mu} > 0$.

Step 5: Convexity in δ. For any μ, the terms $\{p, n\}$, are fixed. The second derivative of δ^a is $(a)(a-1)\delta^a$ which is strictly positive for any $a > 1$.

□

B.3 Related Results Used Elsewhere

In this section, we derived properties that are used later in the text.

Part 5. Inverse productivity. Now, we study the inverse of productivity. Let

$$q(\mu;\delta) = \mathcal{A}^{-1}(\mu;\delta).$$

The function has limits:

$$\lim_{\mu\to 0} q(\mu;\delta) = \delta^{-1}, \qquad \lim_{\mu\to 1} q(\mu;\delta) = \infty, \qquad \lim_{\delta\to 0} q(\mu;\delta) = \infty, \qquad \lim_{\delta\to 1} q(\mu;\delta) = 1.$$

The convexity follows from the concavity of $\mathcal{A}$ and the fact that it takes positive values:

$$q_{\mu\mu} = -\frac{\mathcal{A}_{\mu\mu}}{\mathcal{A}^2} + 2\frac{\mathcal{A}_\mu}{\mathcal{A}^3} > 0.$$

Hence, q is convex in μ.

Part 6. Elasticity of $\mathcal{A}$. A useful object in later derivations is the elasticity of $\mathcal{A}$ with respect to μ:

$$\epsilon_{\mu}^{\mathcal{A}} \equiv \frac{\mathcal{A}_{\mu}(\mu)\,\mu}{\mathcal{A}(\mu)}.$$

Using (24), we obtain:

$$\mathcal{A}_{\mu} = -\frac{1-\delta}{1-\mu}\left(\frac{\delta}{1-\delta\mu}\right)\mu$$

$$\frac{\mathcal{A}_{\mu}(\mu)\,\mu}{\mathcal{A}(\mu)} = -\frac{(1-\delta)\mu}{(1-\delta\mu)(1-\mu)} < 0.$$

Recall the relationship,

$$\frac{\partial}{\partial\mu}[\mathcal{A}(\mu)\,\mu] = \mathcal{A}'(\mu)\,\mu + \mathcal{A}(\mu) = \mathcal{A}(\mu)\,(1+\epsilon_{\mu}^{\mathcal{A}}).$$

We have:

$$1+\epsilon_{\mu}^{\mathcal{A}} = \frac{(1-\mu)(1-\delta\mu)-(1-\delta)\mu}{(1-\delta\mu)(1-\mu)} = \frac{1-\mu+\delta\mu^2-\mu}{(1-\delta\mu)(1-\mu)} = \frac{1}{1-\delta\mu} - \frac{\mu}{1-\mu}.$$

In later proofs, we require obtaining the limits of $1+\epsilon_{\mu}^{\mathcal{A}}$ and establishing monotonicity.

The limits of the function that governs the sign are:

$$\lim_{\mu\to 0}\left[\frac{1}{1-\delta\mu} - \frac{\mu}{1-\mu}\right] = 1 > 0, \text{ and } \lim_{\mu\to 1}\left[\frac{1}{1-\delta\mu} - \frac{\mu}{1-\mu}\right] = -\infty.$$

Since $\epsilon_{\mu}^{\mathcal{A}}$ is continuous in μ, the sign is ambiguous.

Next, to establish monotonicity, notice that

$$\frac{\partial}{\partial\mu}\left[\frac{1}{1-\delta\mu} - \frac{\mu}{1-\mu}\right] = \left[\frac{\delta}{(1-\delta\mu)^2} - \frac{\mu}{1-\mu}\left(\frac{1}{1-\mu} - \frac{1}{\mu}\right)\right] = \frac{\delta}{(1-\delta\mu)^2} - \frac{1}{(1-\mu)^2} < 0.$$

The numerator of the first term is lower than that of the second, whereas the opposite is true about the denominators. Hence, the function is monotone decreasing. Thus, there's a unique crossing point where the function $1+\epsilon_{\mu}^{\mathcal{A}}$ becomes negative.

C Extensions to Other Assignments

C.1 Correlated Assignments

In the main text, we considered payment chains where spot and chained orders have equal probabilities of being linked to orders of the same type. However, our empirical analysis in Table A.4 reveals that payment types exhibit significant correlation: chained orders predominantly connect to other chained orders, while spot orders tend to cluster together. This correlation in order types affects the distribution of chain lengths and ultimately productivity. To examine this, we allow the probability that chained orders link to other chained orders to differ from the probability that spot orders connect to chained orders. We formalize this using a Markov chain representation.

Markov Chain Representation. Consider a payment-chain network where order types (spot or chained) exhibit correlation along production chains. To ease the notation, let $\mu_x = \mu$ denote the fraction of chained orders and $\mu_s = 1 - \mu_x$ the fraction of spot orders. Starting from a spot order, the entire payment network is characterized by a sequence $S_t = \{s, x, s, x, s, \ldots\}$ where $\{s, x\}$ in position i represents whether the i-th order is spot or chained, respectively. If links across chained and spot orders only depend on the current order type, the sequence can be generated by a Markov chain:

$$Q = \begin{bmatrix} p_{ss} & 1 - p_{xx} \\ 1 - p_{ss} & p_{xx} \end{bmatrix} \tag{27}$$

where p_{ss} is the probability that a spot order is linked to a spot, and p_{xx} the corresponding probability that a chained order is linked to a chained order.

For a large N, the stationary distribution must satisfy:

$$\boldsymbol{\mu} = Q\boldsymbol{\mu}, \quad \text{where } \boldsymbol{\mu} = \begin{bmatrix} \mu_s \\ \mu_x \end{bmatrix}. \tag{28}$$

That is, the transition matrix must guarantee that population shares aggregate.

To see how the construction works, we verify the formulas for the i.i.d. case in which transition probabilities equal unconditional probabilities:

$$Q^{\text{iid}} = \begin{bmatrix} \mu_s & \mu_s \\ \mu_x & \mu_x \end{bmatrix}. \tag{29}$$

This trivially satisfies $\boldsymbol{\mu} = Q^{\text{iid}}\boldsymbol{\mu}$ since each column sums to one.

To introduce correlation, we parameterize deviations from independence as:

$$p_{ss} = \mu_s\kappa_s, \quad p_{xx} = \mu_x\kappa_x \tag{30}$$

where $\kappa_s, \kappa_x > 0$ are correlation factors. When the corresponding κ is positive, orders have a greater chance of matching with orders of the same type than under an i.i.d. assignment; when κ is negative, the random assignment overrepresents mixed links.

For positive values of κ, the transition matrix becomes:

$$Q = Q^{\text{iid}} + \tilde{Q} \tag{31}$$

where the deviation matrix is:

$$\tilde{Q} = \begin{bmatrix} \mu_s(\kappa_s - 1) & -\mu_x(\kappa_x - 1) \\ -\mu_s(\kappa_s - 1) & \mu_x(\kappa_x - 1) \end{bmatrix}. \tag{32}$$

There are some parameter constraints: We require a stationarity condition on $\tilde{Q}\boldsymbol{\mu} = \mathbf{0}$ that guarantees the preservation of population shares:

$$\mu_s^2(\kappa_s - 1) = \mu_x^2(\kappa_x - 1). \tag{33}$$

Defining $q \equiv \mu_s(\kappa_s - 1)$ and using the constraint, we obtain:

$$\mu_x(\kappa_x - 1) = \frac{\mu_s}{\mu_x}q. \tag{34}$$

Setting $\rho \equiv \frac{\mu_s}{\mu_x} q$, the transition matrix simplifies to:

$$Q(\rho) = \begin{bmatrix} \mu_s + q & 1 - (\mu_x + \rho) \\ 1 - (\mu_s + q) & \mu_x + \rho \end{bmatrix}. \tag{35}$$

To ensure valid probabilities ($0 \leq p_{ij} \leq 1$), we require: if $\mu_s > 1/2$, then $-\mu_s \leq q \leq 1 - \mu_s$; if $\mu_x > 1/2$, then $-\mu_x \leq \frac{\mu_s}{\mu_x} q \leq 1 - \mu_x$. Combining these constraints:

$$-\min\{\mu_x, \mu_s\} \leq \rho \leq \min\{1 - \mu_s, 1 - \mu_x\}. \tag{36}$$

Chain Length Distribution. Conditional on featuring some chained orders, the length of a chain (consecutive chained orders) follows a geometric distribution. A chain of length n occurs with probability:

$$\Pr(n | n \geq 1) = P(\text{chain ends}) \cdot P(\text{continue chain})^{n-1}. \tag{37}$$

Given the parameterization:

$$P(\text{continue chain}) = p_{xx} = \mu_x + \rho \tag{38}$$

$$P(\text{chain ends}) = 1 - p_{xx} = 1 - \mu_x - \rho \tag{39}$$

Therefore:

$$\Pr(n | n \geq 1) = (1 - \mu_x - \rho)(\mu_x + \rho)^{n-1}. \tag{40}$$

This is a geometric distribution with success probability $p = \mu_x + \rho$.

The expected chain length conditional on some chained orders is:

$$\mathbb{E}[n | n \geq 1] = \frac{1}{1 - p} = \frac{1}{1 - \mu_x - \rho}. \tag{41}$$

Thus, the unconditional distribution is:

$$\mathbb{E}[n] = \mathbb{E}[n | n \geq 1] \Pr[n \geq 1] = \frac{\mu_x - q}{1 - \mu_x - \rho} = \frac{\mu_x - \frac{\mu_x}{\mu_s} \rho}{1 - \mu_x - \rho},$$

where we used that $q \equiv \frac{\mu_x}{\mu_s} \rho$. Notice that when $\rho > 0$, we have positive correlation—chained orders cluster together, increasing average chain length; when $\rho < 0$, the correlation decreases average chain length. When $\rho = 0$ (independence), we recover $\mathbb{E}[L] = \frac{\mu}{1-\mu}$.

Output and TFP. Using the same steps as we do for the i.i.d. case, using the chain-length distribution method, we obtain:

$$\mathcal{A}(\mu; \delta, \rho) = \delta \frac{1 - (\mu_x + \rho)}{1 - \delta(\mu_x + \rho)}. \tag{42}$$

Output is computed by weighting this distribution:

$$\mathcal{Y}(\mu; \delta, \rho) = (1 - \mu) + \mu \mathcal{A} = 1 - \mu \frac{1 - \delta}{1 - \delta(\mu + \rho)}. \tag{43}$$

Again, this verifies the consistency of the formulas when $\rho = 0$.

C.2 Pooled Chained Orders

Next, we consider assignments where chained orders can be efficiently reassigned to chains of approximately equal length, while the mass of spot-to-spot orders remains unchanged. Under this mechanism, the income from every chained order is immediately used to finance another chained order as long as unfunded chained orders exist. This captures scenarios where agents placing multiple chained orders pool their funds, ensuring no cash remains idle while chained orders await funding.

Setup with Correlation. As with correlated assignments, the probability that a spot order is linked to a chained order is $(1-\mu_s-q)$, where q is the correlation parameter from our analysis above. While spot-to-spot matches follow a random process, any spot order that connects to a chained order enters a pool where chains are optimally formed to maximize total production. Thus, the mass of spot orders that can potentially match with chained orders is constant:

$$\bar{N}_s = (1 - \mu)(1 - \mu_s - q)N = (1 - \mu)(\mu - q)N. \tag{44}$$

The ratio of chained orders to available spot orders is:

$$\xi(q) = \frac{N_x}{\bar{N}_s} = \frac{\mu N}{(1 - \mu)(\mu - q)N} = \frac{\mu}{(1 - \mu)(\mu - q)} \tag{45}$$

When $q > 0$ (positive correlation): Fewer spot orders connect to chained orders, increasing ξ; when $q < 0$ (negative correlation): More spot orders connect to chained orders, decreasing ξ toward the efficient ratio. When $q = 0$ (independence), we obtain $\xi(0) = \frac{1}{1-\mu}$, thus the ratio of chained to spot orders exceeds $\frac{\mu}{1-\mu}$. This reveals that the assignment is less efficient than the fully efficient case, since some spot orders are inefficiently matched with other spot orders. When $q = -(1-\mu)$, we have $\xi(-(1-\mu)) = \frac{\mu}{1-\mu}$, recovering the fully efficient assignment case.

Optimal Chain Length Distribution. Following the arguments of the efficient case, the efficient assignment creates chains of two adjacent lengths for chains of length greater than or equal to one:

$$n(q) = \lfloor \xi(q) \rfloor \tag{46}$$

$$n(q) + 1 = \lceil \xi(q) \rceil \tag{47}$$

The fraction of chains with the shorter length is:

$$p(q) = \lceil \xi(q) \rceil - \xi(q) \tag{48}$$

The resulting distribution is:

$$g(0) = 1 - (\mu - q) \quad \text{(spot-only transactions)} \tag{49}$$

$$g(n(q)) = p(q)(\mu - q) \quad \text{(shorter chains)} \tag{50}$$

$$g(n(q) + 1) = (1 - p(q))(\mu - q) \quad \text{(longer chains)} \tag{51}$$

Clearly, the pooling mechanism partially corrects the inefficiency from random matching, but maintains the inefficiencies provoked by the assignment of spot-to-spot links.

Production and TFP. TFP is in this case given by:

$$\mathcal{A}(\mu, q) = \frac{p(q) \cdot Y(n(q)) + (1 - p(q)) \cdot Y(n(q) + 1)}{p(q) \cdot n(q) + (1 - p(q)) \cdot (n(q) + 1)} \tag{52}$$

where $Y(n) = \delta \frac{1-\delta^n}{1-\delta}$. Expected production per order follows the baseline formula, $\mathcal{Y}(\mu, q) = (1 - \mu) + \mu \mathcal{A}(\mu, q)$.

C.3 Endogenous Reallocation

The pooled assignment above assumes costless reallocation of funds across chained orders. However, it is natural to expect that reallocation only occurs for orders in positions exceeding a certain threshold. We now examine a model where chains beyond a threshold length η must reallocate their excess orders to shorter chains, while chains below this threshold operate normally.

Setup. Consider a payment network where chains longer than the threshold η must reallocate their excess orders. We define the **orders needing reallocation (N^r)** as the total number of chained orders in positions beyond η:

$$N^r \equiv \sum_{n=\eta+1}^{\infty} (n-\eta) \cdot g(n) = \frac{\mu^{\eta+1}}{1-\mu} \tag{53}$$

This counts all orders in chains of length $n > \eta$ that exceed the threshold.

The **available slots (N^a)** in other chains is the total unused capacity in chains of length $n \leq \eta$:

$$N^a \equiv \sum_{n=0}^{\eta} (\eta - n) \cdot g(n) = \eta - \left(\frac{\eta\mu^{\eta+1}}{1-\mu}\right)(2-\mu) - \frac{1}{1-\mu} + \frac{(1+\eta)\mu^{\eta}}{1-\mu} \tag{54}$$

Each chain of length $n \leq \eta$ can potentially accept $(\eta - n)$ additional orders before reaching the threshold.

The ratio of reallocating orders to available slots determines the efficient assignment structure:

$$\xi(\eta) = \frac{N^r}{N^a} = \frac{\mu^{\eta}}{(\eta-1)(1-\mu) + \mu^{\eta}[1+\eta(1-\mu)^2]} \tag{55}$$

This ratio determines the optimal chain lengths for efficient assignment:

$$n(\eta) = \lfloor \xi(\eta) \rfloor \tag{56}$$

$$n(\eta) + 1 = \lceil \xi(\eta) \rceil \tag{57}$$

with fraction $p(\eta) = \lceil \xi(\eta) \rceil - \xi(\eta)$ assigned to the shorter length.

In this setting, orders are reallocated sequentially by approaching the customers and suppliers in the last position of an existing chain. They offer to buy from the supplier (corresponding to a spot order) and sell to the last customer order (corresponding to a chain order). When abandoning the chain, the final customer buys from the preceding customer. The reallocation has a cost of Ξ.

Chain-Length Distribution. Because the assignment is sequential, reassigned orders will be placed at the shortest available slots, including those from spot orders. As a result, the distribution will be efficient among the reassigned orders.

Let $g(k) = (1-\mu)\mu^{k-1}$ for $k \geq 1$ denote the baseline geometric distribution for chain lengths, and $G(k) = \sum_{j=1}^{k} g(j) = 1 - \mu^k$ its cumulative distribution function. Note that in the baseline model, spot orders have mass $(1-\mu)$.

After the re-assignment, the resulting hybrid distribution for chain lengths is:

$$\hat{g}(k;\eta,\mu) = \begin{cases} \alpha \cdot p(\eta) & k = n(\eta) \\ \alpha \cdot (1 - p(\eta)) & k = n(\eta) + 1 \\ g(k) & n(\eta) + 2 \leq k \leq \eta - 1 \\ g(\eta) + (1 - G(\eta) & k = \eta \text{ (truncated chains)} \\ 0 & \text{otherwise} \end{cases} \tag{58}$$

where $\alpha = G(n(\eta) + 1)$ is the CDF of chains that would have been length $\leq n(\eta) + 1$ in the baseline. This distribution has three key features: (i) all spot orders have been converted into chains through the efficient assignment, (ii) the original distribution between lengths $n(\eta) + 1$ and $\eta - 1$, is preserved reflecting that orders in those positions are not re-assigned, (iii) the mass at η equals μ^η, capturing that all chains with length above η become chains of length η.

Endogenous Choice. Given the cost Ξ and the sequential nature of the re-assignment, the value of η can be found by solving for the value for which:

$$\Xi < \delta^{n(\eta)+1} - \delta^{\eta+1}, \qquad \Xi \geq \delta^{n(\eta)+1} - \delta^{\eta}.$$

The left inequality ensures reallocation from position $\eta + 1$ to the last position with positive mass $n(\eta) + 1$ is profitable, while the right inequality ensures reallocation from position η is not.

C.4 Default and Cascading

The reassignment model above assumes agents can identify and relocate to shorter chains. However, this requires substantial information about the network. We now consider a simpler response: when chains exceed a threshold η, orders beyond this point are withdrawn rather than reassigned. This captures situations where extended payment chains become unsustainable. Orders in positions $n > \eta$ simply default, keeping their funds. When they do so, all of the subsequent transactions experience a cascading sequence of breakdowns.

Chain Length Distribution. As before, chain length follows a geometric distribution:

$$g(n) = (1-p)p^{n-1} \tag{59}$$

where p generalizes to include correlation effects: $p = \mu + \rho$ in the correlated case.

We capture default by modifying the output function:

$$y_x(n) = \begin{cases} \delta \frac{1-\delta^n}{1-\delta} & \text{if } n \leq \eta \\ \delta \left(\frac{1-\delta^\eta}{1-\delta}\right) & \text{if } n > \eta. \end{cases} \tag{60}$$

The second that transactions above position η get truncated.

Output and TFP. Average production per link in a chain of length n is:

$$\bar{y}_x(n) = \frac{y_x(n)}{n} = \frac{\delta}{n}\frac{1-\delta^n}{1-\delta} - \bar{\gamma}(n) \tag{61}$$

where the default adjustment is:

$$\bar{\gamma}(n) = \begin{cases} 0 & \text{if } n \leq \eta \\ \frac{1}{n}\frac{\delta^\eta}{1-\delta}(1-\delta^{n-\eta}) & \text{if } n > \eta. \end{cases} \tag{62}$$

Using the methods presented for the baseline case, the formula for TFP becomes:

$$\mathcal{A}_\eta = \frac{\sum_{n=1}^{\infty} g(n)\bar{y}_x(n)}{\sum_{n=1}^{\infty} g(n)n}. \tag{63}$$

Combining all terms we obtain that:

$$\mathcal{A}_\eta = \frac{\delta(1-p)}{1-p\delta} - \frac{\delta(1-p)}{1-p\delta}\delta^\eta p^\eta = \mathcal{A}(1-\delta^\eta p^\eta) \tag{64}$$

where $\mathcal{A} = \frac{\delta(1-p)}{1-p\delta}$ is the TFP without default. The truncation factor $(1-\delta^\eta p^\eta)$ captures the productivity loss from default. This loss increases the lower threshold η (more chains are truncated). When $\eta \to \infty$, we recover $\mathcal{A}_\eta \to \mathcal{A}$, the case without default. Output follows the formula for the correlated case, replacing $\mathcal{A}_\eta$ with $\mathcal{A}$.

C.5 Partial Delays

We now consider a hybrid payment system in which each transaction in a chain has a probability q of occurring with a delay. This captures heterogeneity in payment processing—some transactions use sophisticated instant payment systems while others require traditional settlement with delays.

Modified Production Structure. For a chain of length n where k transactions happen, production depends critically on the *location* of delays. We assume the first k links are paid immediately, yielding:

$$y(k,n) = (n-k) + \delta\frac{1-\delta^k}{1-\delta} \tag{65}$$

where $(n-k)$ represents immediate payments and the second term captures delayed payments.

Distribution of Delays and TFP. The number of delayed payments in a chain of length n follows a binomial distribution:

$$M(k,n) = \binom{n}{k} q^k (1-q)^{n-k}. \tag{66}$$

The expected production for a chain of length n is:

$$\mathbb{E}[y(k,n)|n] = \sum_{k=0}^{n} y(k,n) \times M(k,n). \tag{67}$$

To evaluate this expectation, we use three key properties of the binomial distribution:

$$\sum_{k=0}^{n} M(k,n) = 1 \tag{68}$$

$$\sum_{k=0}^{n} k \cdot M(k,n) = qn \tag{69}$$

$$\sum_{k=0}^{n} \delta^k M(k,n) = (1-q+\delta q)^n. \tag{70}$$

The third property follows from the binomial theorem.[58] Substituting these properties:

$$\mathbb{E}[y(k,n)|n] = \sum_{k=0}^{n}(n-k)M(k,n) + \frac{\delta}{1-\delta}\left[1 - \sum_{k=0}^{n} \delta^k M(k,n)\right] \tag{72}$$

$$= (1-q)n + \frac{\delta}{1-\delta}\left[1-(1-q+\delta q)^n\right]. \tag{73}$$

Average production across all chain lengths, weighted by the geometric distribution $g(n) = (1-p)p^{n-1}$ is given by:

$$\mathcal{A}_q = \frac{\sum_{n=1}^{\infty} g(n) \cdot \mathbb{E}[y(k,n)|n]}{\sum_{n=1}^{\infty} g(n) \cdot n} \tag{74}$$

The denominator equals $\frac{1}{1-p}$. For the numerator:

$$(1-p)\sum_{n=1}^{\infty} p^{n-1}\mathbb{E}[y(k,n)|n] \tag{75}$$

$$= \frac{1-q}{1-p} + \frac{\delta}{1-\delta}\left[1-(1-p)\frac{1-q+\delta q}{1-p(1-q+\delta q)}\right] \tag{76}$$

Simplifying yields:

$$\mathcal{A}_q = (1-q) + \delta\frac{q(1-p)}{1-p(1-q+\delta q)} \tag{77}$$

[58] The assumption that delays occur sequentially from the beginning of the chain is crucial. If delays were randomly distributed throughout the chain, the production function would be more complex, as non-delayed segments would be interrupted by delayed ones. The sequential structure ensures that all immediate payments (positions $k+1$ through n) can be realized without interruption. For the third property, recall that

$$\sum_{k=0}^{n}\binom{n}{k}(\delta q)^k(1-q)^{n-k} = (1-q+\delta q)^n. \tag{71}$$

This generalizes our baseline formula: When $q = 1$ (full delays): $\mathcal{A}_q = \frac{\delta(1-p)}{1-p\delta}$. When $q = 0$ (no delays): $\mathcal{A}_q = 1$ When $p = \mu$ and $q = 1$, it recovers the baseline uncorrelated case.

D Proofs for the Partial Equilibrium Analysis (Section 4)

D.1 Entrepreneur's Problem: Details

This appendix provides a detailed characterization of the entrepreneur's problem, including formal statements and proofs of the key results presented in the main text.

Thresholds. The following lemma reproduces Lemma 1 in the main text, which identifies two debt threshold points that determine whether an entrepreneur's expenditure is a mix between spot and chained purchases or whether expenditures are concentrated in one type of order.

Lemma 10. (Expenditure Threshold Points): *Define the* ***efficiency threshold****,* $B^{\star}_{t+1} \equiv R_{t+1}\left(\tilde{B}_t - 1\right)$*. Then:*

(i) $S_{t+1} = 0$ *if and only if* $B_{t+1} > \tilde{B}_{t+1}$*.*

(ii) If $B_{t+1} < \tilde{B}_{t+1}$*,* $X_t > 0$ *if and only if* $B_{t+1} > B^{\star}_{t+1}$*.*

Proof. Part (i): If $B_{t+1} > \tilde{B}_{t+1}$, then $\bar{S}_{t+1} = \max\{\tilde{B}_{t+1} - B_{t+1}, 0\} = 0$. By conditions (12), $S_{t+1} = \min\{\bar{S}_{t+1}, E_{t+1}\} = 0$. Conversely, if $S_{t+1} = 0$, then either $\bar{S}_{t+1} = 0$ or $E_{t+1} = 0$. Since entrepreneurs must have positive endowments at all t consume at all periods, $E_{t+1} > 0$, implying $\bar{S}_{t+1} = 0$, which occurs if and only if $B_{t+1} > \tilde{B}_{t+1}$.

Part (ii): From equation (12), chained expenditures are positive, $X_t > 0$, if and only if expenditures exceed the limit on spot expenditures, $E_t > \bar{S}_t = \max\{\tilde{B}_t - B_t\, 0\}$. Using the budget constraint $E_t = B_{t+1}R^{-1}_{t+1} + 1 - B_t$, this condition becomes:

$$B_{t+1}R^{-1}_{t+1} + 1 - B_t > \tilde{B}_t - B_t, 0$$

Simplifying yields $B_{t+1} > R_{t+1}(\tilde{B}_t - 1) = B^{\star}_{t+1}$. □

As stated in the text, the efficiency threshold $B^{\star}_{t+1}$ marks the debt level above which entrepreneurs must resort to chained purchases in the current period due to insufficient spot borrowing capacity. The threshold $\tilde{B}_{t+1}$ determines whether entrepreneurs lose access to spot purchases entirely in the following period.

Marginal Prices and Optimality Conditions. The presence of two types of purchases—spot and chained—with different prices creates discontinuities in the marginal cost of buying goods. We formalize this through the concept of marginal prices, reproducing the definition in the main text.

Definition 3. (Average and Marginal Prices):

I. *The **average price** per unit of consumption is* $Q_t \equiv E_t/(S_t + X_t)$.

II. *The **marginal expenditure price** at* t *given debt choice* B' *is:*

$$\tilde{q}_t^E (B') \equiv 1 + (q_t - 1)\,\mathbb{I}_{\left[B' \geq B^\star_{t+1}\right]}$$

III. *The **marginal borrowing price** at* $t+1$ *given debt* B' *is:*

$$\tilde{q}_{t+1}^B (B') \equiv 1 + (q_{t+1} - 1)\,\mathbb{I}_{\left[B' > \tilde{B}_{t+1}\right]}$$

The marginal expenditure price $\tilde{q}_t^E(B')$ represents the price at which the entrepreneur purchases goods at the margin in period t. If the choice of future debt exceeds the efficiency threshold, $B' \geq B^\star_{t+1}$, the marginal price of goods equals q_t, the value of chained goods; otherwise, it equals 1 (the spot price). Similarly, $\tilde{q}_{t+1}^B(B')$ captures the price at which saved resources translate into consumption in period $t+1$. If future debt exceeds the SBL ($B' > \tilde{B}_{t+1}$), any savings yield chained purchases at price q_{t+1} because no spot goods can be purchased; otherwise, they yield spot purchases at price 1 because any additional unit of savings allows an additional unit of spot consumption on the margin.

D.2 Proof of Proposition 2 (Entrepreneur's Euler Equation)

We begin by formulating the entrepreneur's problem in sequential form:

Problem 5. *Given B_0 and $\left\{R_{t+1}, \tilde{B}_t\right\}_{t\geq 0}$,*

$$\max_{\{S_t, X_t\}_{t\geq 0}} \sum_{t\geq 0} \beta^t \log(C_t),$$

subject to the budget constraint, $B_t + E_t = R_{t+1}^{-1}B_{t+1} + 1, \forall t \geq 0$, conditions (12) and $C_t = S_t + X_t$, and to the intra- and inter-period constraints, $S_t \leq \bar{S}_t$ and $B_t \leq \bar{B}$.

We derive the entrepreneur's Euler equation as a necessary condition for optimality. The entrepreneur's total expenditures as a function of B_{t+1} are:

$$E_t = E(B_t, B_{t+1}, R_{t+1}) \equiv 1 - B_t + \frac{B_{t+1}}{R_{t+1}}.$$

Given total expenditures, spot expenditures are:

$$S\left(B_t, \tilde{B}_t, B_{t+1}\right) = \min\left\{\max\left\{\tilde{B}_t - B_t, 0\right\}, 1 - B_t + \frac{B_{t+1}}{R_{t+1}}\right\},$$

and chained expenditures are:

$$q_t X_t = E(B_t, B_{t+1}, R_{t+1}) - S\left(B_t, \tilde{B}_t, B_{t+1}\right).$$

The entrepreneur's consumption (total goods obtained) is:

$$C_t = S_t + X_t \quad = \quad S\left(B_t, \tilde{B}_t, B_{t+1}\right) + \frac{E(B_t, B_{t+1}, R_{t+1}) - S\left(B_t, \tilde{B}_t, B_{t+1}\right)}{q_t}.$$

Thus, we can write current consumption as:

$$C_t = C(B_t, B_{t+1}, \tilde{B}_t, q_t, R_{t+1}) \equiv \frac{E(B_t, B_{t+1}, R_{t+1})}{q_t} + \left(1 - \frac{1}{q_t}\right) S\left(B_t, \tilde{B}_t, B_{t+1}\right).$$

The entrepreneur's intertemporal optimization problem can be written entirely in terms of a sequence of debt levels:

$$\max_{\{B_{t+1}\}_{t\geq 0}} \sum_{t\geq 0} \beta^t \log\left(C(B_t, B_{t+1}, \tilde{B}_t, q_t, R_{t+1})\right)$$

with B_0 given.

Consider the choice of B_{t+1}. Define:

$$\Upsilon_t\left(B_{t+1}; B_t, B_{t+2}, \tilde{B}_t, \tilde{B}_{t+1}, R_{t+1}, R_{t+2}, q_t, q_{t+1}\right) \equiv \log\left(C(B_t, B_{t+1}, \tilde{B}_t, q_t, R_{t+1})\right) + \beta \log\left(C(B_{t+1}, B_{t+2}, \tilde{B}_{t+1}, q_{t+1}, R_{t+2})\right).$$

The function Υ_t is the piece that contains the influence of B_{t+1} in the objective. Then, we have that Υ_t features kinks at the threshold points $\left\{B^{\star}_{t+1}, \tilde{B}_{t+1}\right\}$. These kinks arise from the kinks in the functions E and S–reflecting forced switches between spot and chained purchases.

We analyze the optimality conditions by studying the effect of a marginal change in B_{t+1}. Recall the relation between $\tilde{B}_t$ and $B^{\star}_{t+1}$:

$$B^{\star}_{t+1} = R_{t+1}\left(\tilde{B}_t - 1\right).$$

We do so splitting the marginal conditions depending on whether some spot expenditures are available at t:

Case I: $B_t \geq \tilde{B}_t$ (No spot purchases at t). We distinguish subcases depending on whether debt levels permit spot expenditures tomorrow:

I.a. If $B_{t+1} < \tilde{B}_{t+1}$ (spot purchases available at $t+1$):

$$\Upsilon'_t\left(B_{t+1}\right) = \frac{1}{C_t}\frac{1}{q_t R_{t+1}} - \beta\frac{1}{C_{t+1}}.$$

I.b. If $B_{t+1} > \tilde{B}_{t+1}$ (no spot purchases at $t+1$):

$$\Upsilon'_t\left(B_{t+1}\right) = \frac{1}{C_t}\frac{1}{q_t R_{t+1}} - \beta\frac{1}{C_{t+1}}\frac{1}{q_{t+1}}$$

At $B_{t+1} = \tilde{B}_{t+1}$, the derivative features a discontinuity:

$$\lim_{B' \uparrow \tilde{B}_{t+1}} \Upsilon'_t\left(B'\right) < \lim_{B' \downarrow \tilde{B}_{t+1}} \Upsilon'_t\left(B'\right),$$

where the inequality follows from q_{t+1}. The condition implies that the marginal increase in utility is greater to the right $\tilde{B}_{t+1}$. Since a solution at the kink would require:

$$\lim_{B' \uparrow \tilde{B}_{t+1}} \Upsilon_t' (B') > 0 > \lim_{B' \downarrow \tilde{B}_{t+1}} \Upsilon_t' (B'),$$

no solution will feature $B_{t+1} = \tilde{B}_{t+1}$.

Case II: $B_t < \tilde{B}_t$ (Possible spot purchases at t). We distinguish subcases based on the debt level relative to the thresholds:

II.a. If $B_{t+1} < B^{\star}_{t+1}$ (only spot at t, some spot at $t+1$):

$$\Upsilon_t' (B_{t+1}) = \frac{1}{C_t} \cdot \frac{1}{R_{t+1}} - \beta \frac{1}{C_{t+1}}$$

II.b. If $B_{t+1} \in \left(B^{\star}_{t+1}, \tilde{B}_{t+1}\right)$ (mixed at t, spot at $t+1$):

$$\Upsilon_t' (B_{t+1}) = \frac{1}{q_t} \frac{1}{C_t} \cdot \frac{1}{R_{t+1}} - \beta \frac{1}{C_{t+1}}$$

II.c. If $B_{t+1} > \tilde{B}_{t+1}$ (mixed at t, only chained at $t+1$):

$$\Upsilon_t' (B_{t+1}) = \frac{1}{C_t} \cdot \frac{1}{q_t} \frac{1}{R_{t+1}} - \beta \frac{1}{C_{t+1}} \frac{1}{q_{t+1}}$$

At the threshold points we have the following discontinuities:

$$\lim_{B' \uparrow B^{\star}_{t+1}} \Upsilon_t' (B') > \lim_{B' \downarrow B^{\star}_{t+1}} \Upsilon_t' (B'), \quad \lim_{B' \uparrow \tilde{B}_{t+1}} \Upsilon_t' (B') < \lim_{B' \downarrow \tilde{B}_{t+1}} \Upsilon_t' (B') \tag{78}$$

Using the definition of average price $Q_t = E_t / C_t$ and marginal prices, we express Υ_t' in a compact form:

$$\Upsilon_t' (B_{t+1}) = \frac{Q_t}{E_t} \cdot \frac{1}{R_{t+1} \tilde{q}^E_t (B_{t+1})} - \beta \frac{Q_{t+1}}{E_{t+1}} \cdot \frac{1}{\tilde{q}^B_{t+1} (B_{t+1})}$$

Setting $\Upsilon_t' (B_{t+1}) = 0$ at interior solutions yields:

$$\frac{E_{t+1}/E_t}{\beta} \cdot \frac{Q_t}{Q_{t+1}} = \frac{R_{t+1}}{\tilde{q}^B_{t+1}(B_{t+1}) / \tilde{q}^E_t (B_{t+1})}. \tag{79}$$

We already established that $B_{t+1} = \tilde{B}_{t+1}$ is not a possible outcome. Hence, the only relevant corner solution is $B_{t+1} = B^{\star}_{t+1}$. Optimality at that corner solution is given by:

$$\lim_{B' \uparrow B^{\star}_{t+1}} \Upsilon'_t(B') \geq 0 \geq \lim_{B' \downarrow B^{\star}_{t+1}} \Upsilon'_t(B')$$

Opening up the condition translates to:

$$\beta q_t R_{t+1} \geq \frac{E_{t+1}}{E_t} \geq \beta R_{t+1}. \tag{80}$$

Arriving at equations (80) and (79) completes the proof of Proposition 2. □

The Euler equation captures how the entrepreneur's consumption-saving decision depends on both current and future marginal prices, which vary discontinuously with debt levels due to the spot borrowing line constraint.

D.3 Proof of Proposition 3 (Entrepereneur's Stationary Solution)

In the stationary version of the entrepreneur's problem, $R_t = \beta^{-1}$, the SBL is constant, $\tilde{B}_t = \tilde{B}_{ss}$, and the chained-goods price is constant, $q_t = q_{ss}$. Throughout this proof, we assume $\tilde{B}_{ss} > 1$. For convenience, we suppress the arguments of $B^{\star}(1/\beta, \tilde{B}_{ss})$ and denote its steady-state value simply as $B^{\star}_{ss}$.

We establish the following results:

I. If $B_0 \in [0, B^{\star}_{ss}]$, then $B_t = B_0$ for all t.

II. There exists a threshold $B^h > \tilde{B}_{ss}$ such that:

II.a If $B_0 < B^h$, then $B_t \to B^{\star}_{ss}$ in finite time.

II.b If $B_0 > B^h$, then $B_t = B_0$ for all t.

The proof strategy involves: First, establishing two benchmark value functions, $\bar{V}(B)$ and $\underline{V}(B)$, that bound the entrepreneur's value function. Second, showing that for $B_0 \leq B^{\star}_{ss}$, a constant debt policy is optimal and achieves the upper bound $\bar{V}(B)$. Third, proving that for $B_0 \in (B^{\star}_{ss}, \tilde{B}_{ss}]$, optimal deleveraging occurs in finite time with debt converging to $B^{\star}_{ss}$. Finally, establishing the existence and uniqueness of $B^h > \tilde{B}_{ss}$ such that entrepreneurs with debt above B^h maintain constant debt and $V(B) = \underline{V}(B)$.

Value Function Bounds. Consider two auxiliary problems that provide bounds for the entrepreneur's value function. First, the unconstrained entrepreneur problem:

Problem 6 (Upper Bound)**.**

$$\bar{V}(B) = \max_{B' \leq \bar{B}} \ln(E) + \beta \bar{V}(B')$$

subject to: $B + E = 1 + \frac{B'}{R}$.

Second, an entrepreneur problem where only chained goods are consumed:

Problem 7 (Lower Bound)**.**

$$\underline{V}(B) = \max_{B' \leq \bar{B}} \ln(E/q_{ss}) + \beta \underline{V}(B')$$

subject to: $B + E = 1 + \frac{B'}{R}$.

Lemma 11. *The solutions to the benchmark problems are:*

$$\bar{V}(B) = \frac{\ln(1 - (1-\beta)B)}{1-\beta}$$

and

$$\underline{V}(B) = \frac{\ln(1 - (1-\beta)B) - \ln(q_{ss})}{1-\beta}.$$

In both cases, optimal expenditures are $E = 1 - (1-\beta)B$.

Proof. For $\underline{V}$, we guess and verify $\underline{V}(B) = \frac{\ln(1-(1-\beta)B)-\ln(q_{ss})}{1-\beta}$. Substituting into the Bellman equation and taking first-order conditions with respect to B' yields:

$$\frac{1}{1 - B + B'/R}\frac{1}{R} = \frac{\beta}{1-(1-\beta)B'}.$$

Using $\beta = R^{-1}$, we obtain $B = B'$, confirming constant debt. Expenditures are $E = 1 - (1-\beta)B$, verifying the conjecture. Setting $q_{ss} = 1$ yields $\bar{V}(B)$. □

Note that $V(B) \in [\underline{V}(B), \bar{V}(B)]$ since these correspond to more constrained and relaxed versions of the original problem.

Proof of Part I.

Lemma 12. *For $B \leq B^{\star}_{ss}$, $V(B) = \bar{V}(B)$ and $E = 1 - (1-\beta)B$.*

Proof. We verify that setting $E = 1 - (1-\beta)B$ in the original problem yields the upper bound value. Note that:

$$\begin{aligned} E &\leq 1 - (1-\beta)B^{\star}_{ss} \\ &= 1 + \beta B^{\star}_{ss} - B^{\star}_{ss} \\ &= 1 + \beta\beta^{-1}(\tilde{B}_{ss} - 1) - B^{\star}_{ss} \\ &= \tilde{B}_{ss} - B^{\star}_{ss} \\ &\leq \tilde{B}_{ss} - B. \end{aligned}$$

Thus, consuming only through spot purchases is feasible. With $B' = \beta^{-1}(B+E-1) = B$ and consumption $C = S_t + X_t = 1 - (1-\beta)B$ constant across periods, $V(B)$ attains the upper bound. □

Proof of Part II - Preliminary Results. Define the set of debt levels from which setting $B' = B^{\star}_{ss}$ is optimal:

$$\mathcal{B}_0 \equiv \left[B^{\star}_{ss}, B^{\star}_{ss} + (1 - \frac{1}{q_{ss}})(\tilde{B}_{ss} - B^{\star}_{ss}) \right]. \tag{81}$$

Lemma 13. *If $B' = B^{\star}_{ss}$, then $C = \tilde{B}_{ss} - B$.*

Proof. When $B' = B^{\star}_{ss}$, by definition $Q = 1$. Thus:

$$C = 1 + \frac{B^{\star}_{ss}}{R} - B = 1 + \frac{(\tilde{B}_{ss} - 1)}{\beta R} - B = \tilde{B}_{ss} - B,$$

using $B^{\star}_{ss} = 1/\beta(\tilde{B}_{ss} - 1)$. □

Proposition 6. $B' = B^{\star}_{ss}$ *if and only if* $B \in \mathcal{B}_0$.

Proof. We find values of B such that the sub-differential condition holds:

$$\frac{C(B^{\star}_{ss})}{C(B)} \in [1, q_{ss}],$$

where marginal prices are both 1 since $B' \leq B^{\star}_{ss} < \tilde{B}_{ss}$ and $\beta R = 1$. This yields:

$$\frac{\tilde{B}_{ss} - B^{\star}_{ss}}{\tilde{B}_{ss} - B} \geq 1 \iff B \geq B^{\star}_{ss},$$

and

$$\frac{\tilde{B}_{ss} - B^{\star}_{ss}}{\tilde{B}_{ss} - B} \leq q_{ss} \iff B \leq \tilde{B}_{ss} - \frac{1}{q_{ss}}(\tilde{B}_{ss} - B^{\star}_{ss}) < \tilde{B}_{ss}.$$

□

Proposition 7. *Let* $\gamma \equiv \beta / q_{ss}$ *and define:*

$$\Gamma(t) \equiv \frac{1-\gamma^t}{1-\gamma} - q_{ss}^{-t}\frac{1-\beta^t}{1-\beta}, \quad t \geq 0.$$

Let T *be the unique positive integer satisfying* $\Gamma(T) < 1 < \Gamma(T+1)$.

Any decreasing sequence $\{B_{-t}\}$ *for* $t \in \{0, T\}$ *satisfying the entrepreneur's Euler equation with terminal condition* $B_0 \in \mathcal{B}_0$ *satisfies* $B_{-t} \in \mathcal{B}_{-t}$ *where*

$$\mathcal{B}_{-t} \equiv (B^{\ell}_{-t}, B^{\ell}_{-t-1}] \cap [0, \tilde{B}_{ss}],$$

with

$$B^{\ell}_{-t} \equiv B^{\star}_{ss} + (\tilde{B}_{ss} - B^{\star}_{ss})\Gamma(t).$$

Moreover, $\cup_{t=0}^{T} \mathcal{B}_{-t} = (B^{\star}_{ss}, \tilde{B}_{ss}]$.

Proof. For any $B \in (B^{\star}_{ss}, \tilde{B}_{ss}]$, the entrepreneur makes both spot and chained purchases, so the ratio of marginal prices is q_{ss}. Consider a decreasing sequence of debt levels converging to some $B_0 \in \mathcal{B}_0$. For any $B_{-t} \in (B^{\star}_{ss}, \tilde{B}_{ss}]$, the entrepreneur's Euler equation yields:

$$\frac{C_{-t+1}}{C_{-t}} \cdot \frac{Q_{-t}}{Q_{-t+1}} = \beta R \cdot \frac{\tilde{q}^E_{-t}(B_{-t+1})}{\tilde{q}^B_{-t+1}(B_{-t+1})} = 1/q_{ss},$$

where we used that $\beta R = 1$ in steady state and the ratio of marginal prices equals q_{ss} for debt in this range. This gives us the difference equation:

$$C_{-t} = \frac{1}{q_{ss}} C_{-t+1} \text{ for } B_{-t+1} \in (B^{\star}_{ss}, \tilde{B}_{ss}],$$

with terminal condition $C_0 = \tilde{B}_{ss} - B_0$ (from the previous lemma).

Shifting forward the equation, we obtain:

$$C_0 / C_{-t} = q_{ss}^t.$$

Using the budget constraint, consumption at $-t$ is:

$$C_{-t} = \underbrace{\tilde{B}_{ss} - B_{-t}}_{S_{-t}} + \underbrace{\frac{1}{q_{ss}}(1 + \beta B_{-t+1} - \tilde{B}_{ss})}_{X_{-t}}.$$

Combining these expressions:

$$\tilde{B}_{ss} - B_{-t} + q_{ss}^{-1}(1 + \beta B_{-t+1} - \tilde{B}_{ss}) = q_{ss}^{-t}(\tilde{B}_{ss} - B_0).$$

Rearranging:

$$B_{-t} = \tilde{B}_{ss} - q_{ss}^{-t}(\tilde{B}_{ss} - B_0) + q_{ss}^{-1}(1 + \beta B_{-t+1} - \tilde{B}_{ss}).$$

Since $1 + \beta B_{-t+1} - \tilde{B}_{ss} = \beta(B_{-t+1} - B^{\star}_{ss})$ (using the definition of $B^{\star}_{ss}$), we get:

$$B_{-t} = \tilde{B}_{ss} - q_{ss}^{-t}(\tilde{B}_{ss} - B_0) + \frac{\beta}{q_{ss}}(B_{-t+1} - B^{\star}_{ss}).$$

For the boundary sequence with $B_0 = B^{\ell}_0 = B^{\star}_{ss}$, solving this difference equation yields:

$$B^{\ell}_{-t} = B^{\star}_{ss} + (\tilde{B}_{ss} - B^{\star}_{ss}) \sum_{\tau=0}^{t-1} \gamma^{\tau}(1 - q_{ss}^{-t+\tau}),$$

where $\gamma = \beta/q_{ss}$. This sum simplifies to $\Gamma(t)$ as defined.

Since $B^{\ell}_{-1} = B^{\star}_{ss} + (1 - 1/q_{ss})(\tilde{B}_{ss} - B^{\star}_{ss}) = B^{u}_0$ (the upper bound of $\mathcal{B}_0$), the intervals $\mathcal{B}_{-t}$ form a partition of $(B^{\star}_{ss}, \tilde{B}_{ss}]$. The condition $\Gamma(T) < 1 < \Gamma(T+1)$ ensures that T is the last period for which $B^{\ell}_{-T} \leq \tilde{B}_{ss}$. Since $\Gamma(1) = 1 - q_{ss}^{-1} < 1$ and $\lim_{t\to\infty} \Gamma(t) = 1/(1-\gamma) > 1$, with Γ increasing in t, there exists a unique such T. □

Existence and Uniqueness of B^h. Following the analysis for sequences above $\tilde{B}_{ss}$, we establish:

Corollary 3. *There exists a maximal debt level*

$$\underline{B} = \frac{1 - q_{ss}^{-T}(1 - (1-\beta)B^{\star}_{ss})}{1-\beta},$$

such that deleveraging toward $B^{\star}_{ss}$ following the Euler equation is infeasible.

For values above $\underline{B}$, the policy $B' = B$ is the unique solution consistent with the Euler equation. Since $V(\tilde{B}_{ss}) > \underline{V}(\tilde{B}_{ss})$ while $V(\underline{B}) = \underline{V}(\underline{B})$, by continuity there exists $B^h \in [\tilde{B}_{ss}, \underline{B}]$ such that $V(B) = \underline{V}(B)$ for all $B > B^h$.

To establish uniqueness, we show that if $V(B^d) > \underline{V}(B^d)$ for some $B^d > \tilde{B}_{ss}$, then $V(B) > \underline{V}(B)$ for all $B < B^d$. This follows from the decreasing absolute risk aversion property of log utility: starting from any $B < B^d$, the entrepreneur can consume the annuity of the difference $\Delta = B^d - B$ in chained goods and follow the same deleveraging path, yielding strictly higher value than $\underline{V}(B)$.

This completes the proof of Proposition 3. □

E Proofs in the General Equilibrium Analysis - Section 5

E.1 Proof of Proposition 4

Proof. The proof uses the individual optimization creditors together with the market clearing conditions to arrive at a mapping from current debt levels to allocations:

Step 1. From the creditor's log utility and budget constraint $B_{t+1}/R_{t+1} + C_t = B_t$, optimal consumption is:

$$C(B) = (1-\beta)B \tag{82}$$

This is the annuity value of debt holdings.

Step 2. Market clearing for goods requires total expenditures equal total production (normalized to 1):

$$C(B) + E(B) = 1$$

Hence, entrepreneur expenditures are given by the function:

$$E(B) \equiv 1 - C(B) = 1 - (1-\beta)B. \tag{83}$$

Step 3. Given total expenditures $E(B)$ and the constraint that $q \geq 1$ (with equality only when $\mu = 0$), the entrepreneur maximizes goods obtained by maximizing spot purchases subject to the SBL:

$$S(B, \tilde{B}) \equiv \min\{\bar{S}, E(B)\} = \min\{\max\{0, \tilde{B} - B\}, 1 - (1-\beta)B\}$$

Step 4. The value of chained expenditures follows as the residual:

$$\mu(B, \tilde{B}) \equiv E(B) - S(B, \tilde{B}) = 1 - (1-\beta)B - \min\{\max\{0, \tilde{B} - B\}, 1 - (1-\beta)B\}.$$

This yields:

$$\mu(B, \tilde{B}) = \begin{cases} 0 & B < B^{\star}(\tilde{B}), \\ 1 + \beta B - \tilde{B} & B \in [B^{\star}(\tilde{B}), \tilde{B}], \\ 1 - (1-\beta)B & B > \tilde{B}. \end{cases}$$

Step 5. The equilibrium price of chained goods is determined by the network technology:

$$q(B, \tilde{B}) = \mathcal{A}^{-1}(\mu(B, \tilde{B}))$$

where $\mathcal{A}(\cdot)$ is the network efficiency function.

Step 6. The quantity of chained goods obtained by dividing chained expenditures by their price:

$$X(B, \tilde{B}) = \frac{\mu(B, \tilde{B})}{q(B, \tilde{B})} = \mathcal{A}(\mu(B, \tilde{B})).$$

Step 7. The average price per unit of consumption:

$$Q(B, \tilde{B}) = \frac{E(B)}{S(B, \tilde{B}) + X(B, \tilde{B})} = \frac{S(B, \tilde{B}) + q(B, \tilde{B})X(B, \tilde{B})}{S(B, \tilde{B}) + X(B, \tilde{B})}.$$

□

E.2 Analysis: Preliminary Observations

Marginal Expenditure Price as a function of B. The marginal expenditure price depends on both debt and the SBL: $\tilde{q}_t^E = \tilde{q}^E(B_t, \tilde{B}_t)$ where:

$$\tilde{q}^E(B, \tilde{B}) \equiv q(\mu(B, \tilde{B}))\mathbb{I}_{[B \geq B^\star(\tilde{B})]} + (1 - \mathbb{I}_{[B \geq B^\star(\tilde{B})]}).$$

Recall that $\mu(B, \tilde{B})$ is continuous, starts at zero, increases up to $B = \tilde{B}$, then decreases. Since q_t is monotone in μ_t, the marginal expenditure price follows the same pattern.

At the efficiency threshold:

$$\lim_{B \downarrow B^\star(\tilde{B})} \tilde{q}^E(B, \tilde{B}) = \lim_{\mu \downarrow 0} q(\mu) = \delta^{-1},$$

where δ is the network efficiency parameter. Thus $\tilde{q}^E$ has a discontinuity at $B^\star$ since $\lim_{B \uparrow B^\star(\tilde{B})} \tilde{q}^E = 1 \neq \delta^{-1}$.

Average-to-Marginal Price Ratio as function of B. The ratio $Q/\tilde{q}^E$ is crucial for equilibrium dynamics. For $B < B^\star(\tilde{B})$, since $q = 1$, we have $Q/\tilde{q}^E = 1$. For $B \geq \tilde{B}$, also $\tilde{q}^E = Q = q$, so $Q/\tilde{q}^E = 1$.

In the intermediate range $B \in [B^\star(\tilde{B}), \tilde{B}]$:

$$\frac{Q}{\tilde{q}^E} = \frac{1}{q} \cdot \frac{q}{(1 - S/E) + q(S/E)} = \frac{1}{(1 - S/E) + q(\mu(B, \tilde{B}))(S/E)}.$$

At $B = \tilde{B}$, we have $S/E = 0$, so the function is continuous there. However, as $B \downarrow B^\star(\tilde{B})$:

$$\frac{S}{E} \downarrow 1 \text{ and } q \downarrow \delta^{-1} > 1,$$

yielding $Q/\tilde{q}^E = \delta$ at the right limit of $B^\star$ and $Q/\tilde{q}^E = 1$ at the left limit—a discontinuity.

Analysis of the Marginal Borrowing Price. The marginal borrowing price $\tilde{q}^B_{t+1} = \tilde{q}^B(B_{t+1}, \tilde{B}_{t+1})$ where:

$$\tilde{q}^B(B, \tilde{B}) \equiv q(\mu(B, \tilde{B}))\mathbb{I}_{[B \geq \tilde{B}]} + (1 - \mathbb{I}_{[B \geq \tilde{B}]}).$$

Considering the discountinuities in μ:

$$\tilde{q}^B(B, \tilde{B}) = \begin{cases} 1 & B < B^\star(\tilde{B}), \\ 1 & B \in [B^\star(\tilde{B}), \tilde{B}], \\ q & B > \tilde{B}. \end{cases}$$

At $B = \tilde{B}$, the function has a discontinuity:

$$\lim_{B \downarrow \tilde{B}} \tilde{q}^B(B, \tilde{B}) = q(\mu(\tilde{B}, \tilde{B})) = q(1 - (1 - \beta)\tilde{B}) > 1.$$

The ratio $Q/\tilde{q}^B$:

$$\frac{Q}{\tilde{q}^B} = \begin{cases} 1 & B < B^\star(\tilde{B}), \\ Q & B \in [B^\star(\tilde{B}), \tilde{B}], \\ 1 & B > \tilde{B}, \end{cases}$$

where $\lim_{B \downarrow B^\star(\tilde{B})} Q = 1$.

Average Price Elasticity. The elasticity of average price with respect to debt in general equilibrium (after imposing $E = 1 - (1 - \beta)B$) is critical for policy analysis.

Starting from:

$$Q = \frac{1}{\frac{1}{q}(1 - S/E) + S/E},$$

we obtain:

$$\frac{\partial Q}{\partial B} = -Q \cdot \frac{(1 - S/E)\frac{\partial}{\partial B}[1/q] - (1/q - 1)\frac{\partial}{\partial B}[S/E]}{(1/q(1 - S/E) + S/E)}.$$

The derivative of the expenditure share is given by:

$$\frac{\partial}{\partial B}\left[\frac{S}{E}\right] = \begin{cases} 0 & B < B^{\star}(\tilde{B}) \\ \left[\frac{\tilde{B}-B}{1-(1-\beta)B}\right]\left(\frac{(1-\beta)}{1-(1-\beta)B} - \frac{1}{\tilde{B}-B}\right) & B \in [B^{\star}(\tilde{B}), \tilde{B}) \\ 0 & B \geq \tilde{B}. \end{cases}$$

This derivative is negative in the intermediate region and has discontinuities at the boundaries:

$$\left.\frac{\partial}{\partial B}\left[\frac{S}{E}\right]\right|_{B\uparrow\tilde{B}} = -\frac{1}{1-(1-\beta)\tilde{B}} < -1,$$

and

$$\left.\frac{\partial}{\partial B}\left[\frac{S}{E}\right]\right|_{B\downarrow B^{\star}} = -1.$$

For the price derivative:

$$\frac{\partial}{\partial B}[q] = \frac{\partial q}{\partial \mu} \cdot \frac{\partial \mu}{\partial B} = q_{\mu}\mu\frac{\frac{\partial E}{\partial B} - \frac{\partial S}{\partial B}}{E},$$

where ϵ^{q}_{μ} denotes the elasticity of q with respect to μ. These elasticities determine the general equilibrium feedback effects between debt accumulation and payment chain formation.

E.3 Derivation of the map $\mathcal{B}$

For convenience, we reproduce the aggregate Euler equation (13) here:

$$\underbrace{\frac{B}{1-(1-\beta)B} \cdot Q\left(B, \tilde{B}\right)}_{\equiv \mathcal{E}\left(B;\tilde{B}\right)} = \underbrace{\frac{B'}{1-(1-\beta)B'}\frac{Q\left(B', \tilde{B}'\right)}{\Pi\left(B'; B, \tilde{B}, \tilde{B}'\right)}}_{\equiv \mathcal{E}'\left(B'; B, \tilde{B}, \tilde{B}'\right)}.$$

First, we show that for a subset of $B > \tilde{B}'$ there are two roots B' that solve $\mathcal{E} = \mathcal{E}'$. The roots satisfy $B'_1 < \tilde{B}' < B'_2$ and $B'_2 = B$ thus $R = \beta^{-1}$ and $q = q'$. We then show for any $B \in \left(\tilde{B}', B^h\right)$ the larger root B'_2 cannot be an individual optimum during a transition.

The first result is summarized by the following Lemma.

Lemma 14. *There exists a threshold B^* such that for any $B \in [\tilde{B}', B^*]$, the equation $\mathcal{E} = \mathcal{E}'$ has two roots B': one root is $B = B'$ and the other satisfies $B' \leq \tilde{B}'$. For any $B > B^*$, only $B = B'$ is a solution. B^* solves:*

$$\mathcal{E}'\left(B^*; \tilde{B}, \tilde{B}, \tilde{B}'\right) = \lim_{B' \uparrow \tilde{B}'} \mathcal{E}'\left(B'; \tilde{B}, \tilde{B}, \tilde{B}'\right).$$

The interpretation of B^* is obtained from the left panel of Figure 9: B^* is the highest value for which the Euler equation has a solution $B' \in (\tilde{B}', B^*)$.

Proof. Since $\tilde{B}' > \tilde{B}$, if $B' = B > \tilde{B}'$, the worker only spends chained expenditures. Thus,

$$\Pi\left(B'; B, \tilde{B}, \tilde{B}'\right) = 1 \text{ and } Q\left(B, \tilde{B}\right) = Q\left(B', \tilde{B}'\right) = q\left(1 - (1-\beta) B\right).$$

Hence, $B' = B$ is always a solution to the Euler equation when $B' = B > \tilde{B}'$.

Let $B > B^*$. The function $\mathcal{E}(B)$ is monotone increasing. In turn, the function $\mathcal{E}'(B')$ is monotone increasing to the left of $\tilde{B}'$. Hence, for $B > B^*$ a solution must lie at $B' > \tilde{B}'$. However, for any $B' > \tilde{B}'$,

$$\mathcal{E}'\left(B'; \tilde{B}, \tilde{B}, \tilde{B}'\right) = \frac{B'}{1-(1-\beta) B'} q\left(1-(1-\beta) B'\right)$$

which follows from the fact that $B' > \tilde{B}' > B^{\star}\left(1/\beta, \tilde{B}'\right)$ in which case $\Pi\left(B'; B, \tilde{B}, \tilde{B}'\right) = 1$. This implies that for any $B > B^*$ we have that:

$$\mathcal{E}\left(B; \tilde{B}\right) = \mathcal{E}'\left(B'; B, \tilde{B}, \tilde{B}'\right),$$

which implies,

$$\frac{B}{1-(1-\beta) B} \cdot q\left(1-(1-\beta) B\right) = \frac{B'}{1-(1-\beta) B'} \cdot q\left(1-(1-\beta) B'\right).$$

By monotonicity, we immediately conclude that $B = B'$ and that is the only solution.

Now consider $B \in \left(\tilde{B}', B^*\right)$. We now look for a solution $B' \in \left(B^{\star}(\tilde{B}'), \tilde{B}'\right)$. If $B <$

B^*, again by monotonicity:

$$\mathcal{E}'\left(B;\tilde{B},\tilde{B},\tilde{B}'\right) < \lim_{B'\uparrow\tilde{B}'}\mathcal{E}'\left(B';\tilde{B},\tilde{B},\tilde{B}'\right).$$

We make use of the following calculations:

$$\begin{aligned}\mathcal{E}'\left(B';\tilde{B},\tilde{B},\tilde{B}'\right) = \lim_{B'\downarrow\tilde{B}'}\mathcal{E}'\left(B';\tilde{B},\tilde{B},\tilde{B}'\right) \Rightarrow \\ &= \frac{B'}{1-(1-\beta)B'}\cdot\frac{q\left(1-(1-\beta)\tilde{B}'\right)}{\lim_{B'\uparrow\tilde{B}'}\Pi\left(B';\tilde{B},\tilde{B},\tilde{B}'\right)} \\ &= \frac{\tilde{B}'}{1-(1-\beta)\tilde{B}'}\cdot\frac{q\left(1-(1-\beta)\tilde{B}'\right)}{\frac{q\left(1-(1-\beta)\tilde{B}'\right)}{q\left(1-(1-\beta)\tilde{B}\right)}} \\ &= \frac{\tilde{B}'}{1-(1-\beta)\tilde{B}'}\cdot q\left(1-(1-\beta)\tilde{B}\right) > 0.\end{aligned}$$

Fix $\tilde{B},\tilde{B}'$. Define $B^* > \tilde{B}$ to be the value that solves:

$$\mathcal{E}'\left(B^*;\tilde{B},\tilde{B},\tilde{B}'\right) = \lim_{B'\uparrow\tilde{B}'}\mathcal{E}'\left(B';\tilde{B},\tilde{B},\tilde{B}'\right).$$

By definition, if indeed $B^* > \tilde{B}$, then

$$\frac{B^*}{1-(1-\beta)B^*}\cdot q\left(1-(1-\beta)B^*\right) = \lim_{B'\uparrow\tilde{B}'}\frac{B'}{1-(1-\beta)B'}\cdot\frac{Q\left(B',\tilde{B}'\right)}{\Pi\left(B';B^*,\tilde{B},\tilde{B}'\right)}$$

which implies

$$B^* = \frac{\tilde{B}'}{\frac{1-(1-\beta)\tilde{B}'}{q\left(1-(1-\beta)\tilde{B}'\right)}+(1-\beta)\tilde{B}'} = \frac{\tilde{B}'}{C^w\left(\tilde{B}',\tilde{B}'\right)+C^s\left(\tilde{B}'\right)} > \tilde{B}'.$$

The last inequality verifies that $B^* > \tilde{B}'$. This result follows from

$$\lim_{B'\uparrow\tilde{B}'}\Pi\left(B';B^*,\tilde{B},\tilde{B}'\right) = \left(q\left(1-(1-\beta)B^*\right)\right)^{-1}$$

and

$$\lim_{B'\uparrow\tilde{B}'} Q\left(B', \tilde{B}'\right) = q\left(1-(1-\beta)\,\tilde{B}'\right).$$

The inequality holds since $q > 1$. Furthermore, this satisfies $B^* > \tilde{B}$ as required since the SBL sequence is weakly increasing.[59]

Next, we show that for any $B \in \left(\tilde{B}', B^*\right)$, the equation $\mathcal{E} = \mathcal{E}'$ has two roots; one above $\tilde{B}'$ and one below. The solution above $\tilde{B}'$ is trivial: $B = B' > \tilde{B}' \geq \tilde{B}$ satisfies the equation. For the solution below, we use continuity and monotonicity to show that there is a unique $B' < \tilde{B}'$ such that

$$\begin{aligned}
\mathcal{E}'\left(B'; \tilde{B}, \tilde{B}, \tilde{B}'\right) &= \lim_{B'\downarrow\tilde{B}'} \mathcal{E}'\left(B'; \tilde{B}, \tilde{B}, \tilde{B}'\right) \\
\frac{B'}{1-(1-\beta)\,B'} \cdot Q\left(B', \tilde{B}'\right) &= \frac{B'}{1-(1-\beta)\,B'} \cdot \frac{q\left(1-(1-\beta)\,\tilde{B}'\right)}{\lim_{B'\uparrow\tilde{B}'} \Pi\left(B'; \tilde{B}, \tilde{B}, \tilde{B}'\right)} \\
\frac{B'}{1-(1-\beta)\,B'} \cdot Q\left(B', \tilde{B}'\right) &= \frac{\tilde{B}'}{1-(1-\beta)\,\tilde{B}'} \cdot \frac{q\left(1-(1-\beta)\,\tilde{B}'\right)}{\frac{q\left(1-(1-\beta)\tilde{B}'\right)}{q\left(1-(1-\beta)\tilde{B}\right)}} \\
\frac{B'}{1-(1-\beta)\,B'} \cdot Q\left(B', \tilde{B}'\right) &= \frac{\tilde{B}'}{1-(1-\beta)\,\tilde{B}'} \cdot q\left(1-(1-\beta)\,\tilde{B}\right) > 0
\end{aligned}$$

Lets call this B', $\underline{B}$ and think of the interval $\left(\tilde{B}', B^*\right)$. The interpretation of $\underline{B}$ is that it is the small root for the debt level "$\tilde{B}' + \varepsilon$". Meaning, if $\tilde{B}'$ is the small root for B^*(the end of the interval) then $\underline{B}$ is the small root for the start of the interval. Now, notice that the RHS is a constant with respect to $\tilde{B}'$. To establish existence of $\underline{B}$ further notice that the LHS tends to $\lim_{B'\uparrow\tilde{B}'} \mathcal{E}'\left(B'; B^*, \tilde{B}, \tilde{B}'\right)$ as $B' \uparrow \tilde{B}'$ (this was shown in the previous step) and this is larger than the RHS just by comparing magnitudes ($\tilde{B}' \geq \tilde{B}$ and $q\,(1-(1-\beta)\,B^*) > 1$). Then notice that as $B' \downarrow 0$ the LHS goes to zero which is lower than the RHS. By continuity of the LHS we can apply the intermediate value theorem for existence. Uniqueness is granted since the LHS is increasing in B'. The fraction is clearly increasing in B', the average price too because as $B' \uparrow \tilde{B}'$ the share of chained expenditure increases and also its price does so. This statement also uses the fact that Q is the weighted harmonic mean of prices with expenditure weights. Since the LHS is

[59]This is important because the relevant points to study are those for which the all chained equilibrium is possible. These are the cases where $B^* > \tilde{B}' \geq \tilde{B}$.

continuous and increasing it maps the interval $\left(\underline{B}, \tilde{B}'\right)$ onto

$$\left(\lim_{B'\downarrow\tilde{B}'} \mathcal{E}'\left(B'; \tilde{B}, \tilde{B}, \tilde{B}'\right), \lim_{B'\uparrow B^*} \mathcal{E}'\left(B'; \tilde{B}, \tilde{B}, \tilde{B}'\right)\right).$$

So (since $\mathcal{E}'$ is increasing and continuous) it covers all the image of $\left(\tilde{B}', B^*\right)$. This proves that for $B \in \left(\tilde{B}', B^*\right)$ there exist two roots that solve $\mathcal{E} = \mathcal{E}'$. One is $B_2' = B > \tilde{B}'$ and the other is a $B_1' \in \left(\underline{B}, \tilde{B}'\right)$.

To make affairs clearer, suppose that $B^* < B^h$ then for all $B \in \left(\tilde{B}', B^*\right)$ the larger root $B' = B$ is not an equilibrium and there is a region (B^*, B^h) that does not have a symmetric competitive equilibrium because the only root that solves equation (13)provides allocations and prices that are not an equilibrium. Now suppose that $B^* > B^h$, then for $B \in \left(\tilde{B}', B^h\right)$ the larger root $B' = B$ is not an equilibrium and for $B \in (B^h, \overline{B})$. Summarizing both cases. For $B_0 < B^h\left(\tilde{B}_0\right)$, if a (symmetric competitive) equilibrium exists, then it is given by the smaller root of equation (13).

□

E.4 Proof of Corollary 1

Part i. At $B_t \leq B^\star\left(\beta^{-1}, \tilde{B}_{ss}\right)$ if $R = \beta^{-1}$ then marginal and average prices are equal to 1. Then it is evident that $B_{t+1} = B_t$ solves equation (13) and the expenditure is $1 - (1-\beta) B_t \leq \tilde{B}_{ss} - B_t$ by assumption. As a consequence, the steady state is non-disrupted.

Part ii. This case follows immediately from Proposition 3. If $B_0 > B^h$, there exists an equilibrium with $R_t = 1/\beta$ and $B_t = B_0$. Recall that if $R_t = 1/\beta$, and $B_t = B_{t+1}$ is a solution to $\mathcal{E}_t = \mathcal{E}_{t+1}$. Since for $R_t = 1/\beta$, $B_0 = B_t$ is a solution to the worker's problem, both Euler equations are satisfied and the asset clearing condition holds, sequences with $R_t = 1/\beta$ and $B_t = B_0$ are consistent with the equilibrium definition. Since at $B_t > B^h > B^\star$, $q_t > 1$. Hence, the economy remains permanently disrupted.

E.5 Proof of Corollary 2

That $B_{t+1} < B_t$ if $B_t \in \left(\tilde{B}, B^* \left(\tilde{B}\right)\right)$ is immediate from Lemma 14 since we are choosing the smaller root and the larger root is $B_{t+1} = B_t$. For $B_t > B^\star$ it is enough to show that $\beta R_{t+1} < 1$ as done in step 3 of the proof of Lemma 14.

F Constrained Efficiency Analysis - Section 5

F.1 Characterization of the Primal Problem

This appendix provides the complete characterization of the Primal Planner Problem in subsection 5.3. For convenience, we reproduce the problem here and then derive the solution across different regimes of the spot borrowing limit $\tilde{B}$.

Since the maximization is static (time-separable), we can solve the problem period by period. The Primal Planner's problem at each date is:

Problem 8. (Static Primal Problem):

$$\mathcal{P}^{\theta}\left(\tilde{B}\right)=\max_{B\in\left[0,\bar{B}\right]}\mathcal{P}\left(B,\tilde{B}\right)$$

where

$$\mathcal{P}\left(B,\tilde{B}\right)=\left(1-\theta\right)\log\left(\left(1-\beta\right)B\right)+\theta\log\left(\mathcal{A}\left(\mu\left(B,\tilde{B}\right)\right)\mu\left(B,\tilde{B}\right)+S\left(B,\tilde{B}\right)\right)$$

subject to:

$$\mu\left(B,\tilde{B}\right)\equiv 1-\left(1-\beta\right)B-\min\left\{\max\left\{\tilde{B}-B,0\right\},1-\left(1-\beta\right)B\right\}$$

$$S\left(B,\tilde{B}\right)=\min\left\{\max\left\{\tilde{B}-B,0\right\},1-\left(1-\beta\right)B\right\}.$$

This is a problem in which the planner chooses debt for agents, while respecting the SBL and payment constraints. Then, the optimal path of debt for the planner's problem is $B_t=\mathcal{B}^{\theta}\left(\tilde{B}_t\right)$. We proceed to solve for $\mathcal{B}^{\theta}$. Welfare depends on whether and how the planner may want to distort TFP to increase egalitarian objectives, given its Pareto weight. The tradeoffs depend on the SBL, $\tilde{B}$.

We make use of two auxiliary problems, the best and worst value problems:

$$\underline{\mathcal{P}}^{\theta}\equiv\mathcal{P}^{\theta}\left(0\right)\text{ and }\bar{\mathcal{P}}^{\theta}\equiv\mathcal{P}^{\theta}\left(\bar{B}\right).$$

That, is $\underline{\mathcal{P}}^{\theta}$ is the value corresponding to no SBL constraints and $\bar{\mathcal{P}}^{\theta}$ the corresponding problem in the case where the entrepreneur must only make chained expenditures.

Recall that the unconstrained solution to the Primal Planner's problem is given by the ratio of Pareto weights:

$$\frac{1-(1-\beta)B^o}{(1-\beta)B^o}=\frac{\theta}{1-\theta}\rightarrow B^o=\frac{1-\theta}{1-\beta}.$$

The solution is as follows:

Proposition 8. (Solution of the Primal Problem): *Let the solution to the primal problem be $\mathcal{B}^\theta$. The function $\mathcal{B}^\theta$ satisfies:*

I. Efficiency Threshold. *For $\tilde{B}\geq\frac{1-\theta\beta}{1-\beta}$, $\mathcal{B}^\theta=B^o$. Moreover, for this debt level $X=0$.*

II. Inefficiency Threshold. *For $\tilde{B}<\frac{1-\theta\beta}{1-\beta}$ the planner's solution induces TFP losses. The solution to the Primal Planner's problem in this region depends on the threshold SBL, $\tilde{B}^i$.*

II.a Social Insurance and Productive Efficiency. *For $\tilde{B}\in\left[\tilde{B}^i,\frac{1-\theta\beta}{1-\beta}\right]$, we have that $\mathcal{B}^\theta=B^\star\left(\tilde{B}\right)$:*

$$\frac{1-(1-\beta)B^\star\left(\tilde{B}\right)}{(1-\beta)B^\star\left(\tilde{B}\right)}\leq\frac{\theta}{1-\theta}\frac{1-\beta\delta}{1-\beta}$$

II.b Social Insurance Complements Productive Efficiency. *If $\tilde{B}^s<\tilde{B}^i$, for $\tilde{B}\in\left[\tilde{B}^s,\tilde{B}^i\right]$, we have that $\mathcal{B}^\theta$ solves:*

$$\frac{1-(1-\beta)B^\theta}{(1-\beta)B^\theta}=\frac{\theta}{1-\theta}\frac{Q\left(B^\theta,\tilde{B}\right)}{q\left(B^\theta,\tilde{B}\right)}\left(\frac{q\left(B^\theta,\tilde{B}\right)-\beta\left(1+\epsilon_\mu^{\mathcal{A}}\left(\mu\left(B^\theta,\tilde{B}\right)\right)\right)}{1-\beta}\right).$$

Moreover, for this debt level $X,S>0$.

II.c Social Insurance Conflicts Productive Efficiency. *For $\tilde{B}\in\left[0,\tilde{B}^s\right]$, we have that $\mathcal{B}^\theta$ is the unique constant solution $B^\theta>B_{ss}$ to the equation:*

$$\frac{1-(1-\beta)B^\theta}{(1-\beta)B^\theta}=\frac{\theta}{1-\theta}\left(1+\epsilon_\mu^{\mathcal{A}}\left(1-(1-\beta)B^\theta\right)\right).$$

Moreover, for this debt level $S=0$.

Threshold value $\tilde{B}^i$. *$\tilde{B}^i$ is the unique solution to:*

$$\frac{1-(1-\beta)\,B^{\star}\left(\tilde{B}^i\right)}{(1-\beta)\,B^{\star}\left(\tilde{B}^i\right)} = \frac{\theta}{1-\theta}\frac{1-\beta\delta}{1-\beta}$$

Threshold value $\tilde{B}^s$. *Let $\underline{\mathcal{P}}^{\theta} \equiv \mathcal{P}^{\theta}(0)$ and $\bar{\mathcal{P}}^{\theta} \equiv \mathcal{P}^{\theta}\left(\bar{B}\right)$. The threshold $\tilde{B}^s$ solves:*

$$\bar{\mathcal{P}}^{\theta} = \underline{\mathcal{P}}^{\theta} + \int_{\tilde{B}^s}^{\frac{1-\theta\beta}{1-\beta}} \left(\mathcal{P}_{\tilde{B}} + \mathcal{P}_B|_{B=B^{\star}(\tilde{B})} \cdot B^{\star}_{\tilde{B}}\left(\tilde{B}\right)\right) d\tilde{B}.$$

where for $\tilde{B} \in \left[\tilde{B}^i, \frac{1-\theta\beta}{1-\beta}\right]$ we have that $\mathcal{P}^{\theta}_{\tilde{B}}\left(\tilde{B}\right)$ equals:

$$\frac{\theta}{1-(1-\beta)\,B^p(\tilde{B})}\frac{Q\left(B^p(\tilde{B}),\tilde{B}\right)}{q\left(B^p(\tilde{B}),\tilde{B}\right)}\left(\mathcal{A}\left(\mu\left(B^p(\tilde{B}),\tilde{B}\right)\right)-\left(1+\epsilon^{\mathcal{A}}_{\mu}\left(\mu\left(B^p(\tilde{B}),\tilde{B}\right)\right)\right)\right)$$

and

$$\mathcal{P}_B|_{B=B^{\star}(\tilde{B})} = \frac{1-\theta}{(1-\beta)\,B^{\star}\left(\tilde{B}\right)} - \frac{\theta\,(1-\beta\delta)}{1-(1-\beta)\,B^{\star}\left(\tilde{B}\right)}$$

F.2 Proof of Proposition Solution to Primal Problem

CaseI. Values of $\tilde{B}$ such that all consumption is spot. Ideally, the planner wants to maximize spot consumption by setting $\mu = 0$. The unconstrained solution to the Primal Planner's problem sets $B = B^o$ and yields the value as $\bar{\mathcal{P}}^{\theta}$. With constraints, this value is achieved if and only if all spot consumption is feasible:

$$\max\left\{\tilde{B} - B^o, 0\right\} \geq 1 - (1-\beta)\,B^o > 0.$$

Thus, we need

$$\tilde{B} \geq B^o$$

and that:

$$\tilde{B} \geq 1 + \beta B^o.$$

Combining both constraints, we have:

$$\tilde{B} \geq B^o + \max\left\{1 - (1-\beta) B^o, 0\right\}.$$

We know that the optimal debt $B^o < (1-\beta)^{-1}$, i.e., it must be less than the natural borrowing limit. Hence, we get rid of the max operator:

$$\tilde{B} \geq 1 + \beta B^o = 1 + \frac{\beta}{1-\beta}(1-\theta) = \frac{1-\theta\beta}{1-\beta}. \tag{84}$$

Thus, for these levels of the SBL that satisfy the condition above, the planner can achieve the unconstrained first best.

Case 2. Values of $\tilde{B}$ such that some consumption is chained. Now consider the case where the constraint binds, $\tilde{B} < \frac{1-\theta\beta}{1-\beta}$. In this case, the planner cannot achieve the unconstrained solution. The amount of chained expenditures are therefore positive:

$$\mu\left(B, \tilde{B}\right) = 1 - (1-\beta) B - \min\left\{\max\left\{\tilde{B} - B, 0\right\}, 1 - (1-\beta) B\right\} > 0.$$

We have critical values.

- If the planner chooses $B^p \geq \tilde{B}$, there is no spot consumption.
- If the planner chooses $B^p < B_l\left(\tilde{B}\right) < \tilde{B}$, there is no chained consumption. The threshold $B_l\left(\tilde{B}\right)$ solves,

  $$\max\left\{\tilde{B} - B_l, 0\right\} = 1 - (1-\beta) B_l.$$

  Since $B_l < \tilde{B}$, we obtain:

  $$\tilde{B} = 1 + \beta B_l \rightarrow B_l\left(\tilde{B}\right) = \max\left\{0, \beta^{-1}\left(\tilde{B} - 1\right)\right\} = B^\star\left(\tilde{B}\right).$$

  $B^\star\left(\tilde{B}\right) < B^o$ since we are in the constrained region.

Consider now the planner problem that restricts choices to at least some of both goods are consumed by the entrepreneur:

Problem 9. *The Primal Planner's problem restricted to both types of consumption is:*

$$\tilde{\mathcal{P}}\left(\tilde{B}\right)=\max_{B}\left\{(1-\theta)\log\left((1-\beta)B\right)+\theta\log\left(\mathcal{A}\left(\mu\left(B,\tilde{B}\right)\right)\cdot\mu\left(B,\tilde{B}\right)+S\left(B,\tilde{B}\right)\right)\right\}$$

subject to:

$$B\in\left[B^{\star}(\tilde{B}),\tilde{B}\right]$$

$$\mu\left(B,\tilde{B}\right)\equiv 1+\beta B-\tilde{B}$$

and

$$S\left(B,\tilde{B}\right)=\tilde{B}-B.$$

We have the following Lemma.

Lemma 15. *For any $\tilde{B}<\frac{1-\theta\beta}{1-\beta}$, the original planner problem satisfies:*

$$\mathcal{P}^{\theta}\left(\tilde{B}\right)=\max\left\{\tilde{\mathcal{P}}\left(\tilde{B}\right),\underline{\mathcal{P}}^{\theta}\right\}.$$

Proof. Indeed, in the region $B^{p}\in\left[0,B^{\star}(\tilde{B})\right]$ the objective of the planner is equivalent to the objective when the SBL is most relaxed, $\mathcal{P}\left(B,\bar{B}\right)$. Thus, since $\tilde{B}<\frac{1-\theta\beta}{1-\beta}$, the planner's objective is increasing in the region $\left[0,B_{l}\left(\tilde{B}\right)\right]$. Thus, the planner's solution must fall in between $B^{p}\in\left[B_{l}\left(\tilde{B}\right),\bar{B}\right]$. For any $B^{p}\geq B_{h}\left(\tilde{B}\right)=\tilde{B}$, the objective function in $\mathcal{P}\left(B,\tilde{B}\right)$ is independent of $\tilde{B}$ and hence, must coincide with the value of $\underline{\mathcal{P}}^{\theta}$. Hence, we can partition $\mathcal{P}^{\theta}\left(\tilde{B}\right)$ according to the Lemma. □

To prove the main result, we solve the problems $\underline{\mathcal{P}}^{\theta}$ and $\tilde{\mathcal{P}}\left(\tilde{B}\right)$.

Auxiliary Problem $\underline{\mathcal{P}}^{\theta}$: no spot consumption. The planner's problem with the tightest SBL $\underline{\mathcal{P}}^{\theta}=\mathcal{P}^{\theta}(0)$ is given by:

$$\underline{\mathcal{P}}^{\theta}=\max_{B\in\left[0,\bar{B}\right]}\underline{\mathcal{P}}\left(B,0\right)$$

where

$$\underline{\mathcal{P}}^{\theta}=\max_{B\in\left[0,\bar{B}\right]}\left\{(1-\theta)\log\left((1-\beta)B\right)+\theta\log\left(\mathcal{A}\left(\mu\left(B,0\right)\right)\mu\left(B,0\right)\right)\right\}$$

subject to:

$$\mu\left(B,0\right)\equiv 1-(1-\beta)B.$$

To solve this problem, we perform some calculations. First, note that:

$$\frac{\partial\left[\mathcal{A}\left(\mu\right)\mu\right]}{\partial\mu}=\mathcal{A}\left(\mu\right)\left(1+\epsilon^{\mathcal{A}}\right)$$

where,

$$\epsilon_{\mu}^{\mathcal{A}}\equiv\frac{\partial\mathcal{A}\left(\mu\right)}{\partial\mu}\frac{\mu}{\mathcal{A}\left(\mu\right)}.$$

The derivative $\underline{\mathcal{P}}_{B}$ is therefore given by:

$$\underline{\mathcal{P}}_{B}=\left(1-\theta\right)\frac{\left(1-\beta\right)}{\left(1-\beta\right)B}+\theta\frac{\mathcal{A}\left(\mu\right)\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\mu_{B}\left(B,0\right)}{\mathcal{A}\left(\mu\right)\mu}=\frac{\left(1-\theta\right)}{B}-\theta\frac{\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\left(1-\beta\right)}{1-\left(1-\beta\right)B}.$$

The second equation uses that $\mu_{B}\left(B,0\right)\equiv-\left(1-\beta\right).$

The first term in $\underline{\mathcal{P}}_{B}$, $\left(1-\theta\right)/B$, is decreasing in B. The second term,

$$\theta\frac{\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\left(1-\beta\right)}{1-\left(1-\beta\right)B},\tag{85}$$

is increasing. We know this because the denominator is decreasing in $B\in\left[0,\bar{B}\right]$ and the elasticity of TFP is itself decreasing in μ,

$$\epsilon_{\mu\mu}^{\mathcal{A}}=\frac{\partial}{\partial\mu}\left[\frac{\mathcal{A}'\left(\mu\right)\mu}{\mathcal{A}\left(\mu\right)}\right]=\frac{\mathcal{A}''\left(\mu\right)\mu}{\mathcal{A}\left(\mu\right)}+\frac{\mathcal{A}'\left(\mu\right)}{\mathcal{A}\left(\mu\right)}-\frac{\left[\mathcal{A}'\left(\mu\right)\right]^{2}}{\left[\mathcal{A}\left(\mu\right)\right]^{2}}\mu<0.$$

Hence, $\epsilon_{\mu\mu}^{\mathcal{A}}\mu_{B}\left(B,0\right)>0$, since the product of two negative numbers is positive, thus the numerator of the second term (85) is increasing in B. Thus, $\underline{\mathcal{P}}$ is concave and therefore $\underline{\mathcal{P}}^{\theta}$ has a unique solution:

$$\frac{1-\left(1-\beta\right)B}{\left(1-\beta\right)B}=\frac{\theta}{1-\theta}\left(1+\epsilon_{\mu}^{\mathcal{A}}\left(\mu\left(B,0\right)\right)\right)$$

and recall that $\mu\left(B,0\right)=1-\left(1-\beta\right)B$. We call this solution $\underline{B}^{p}$: the planner debt level under the most tight SBL. We have the following Lemma:

Lemma 16. *The solution* $\underline{B}^{p}>B^{o}$.

Proof. The proof is immediate from $1+\epsilon_{\mu}^{\mathcal{A}}<1$ and the fact that

$$\frac{1-\left(1-\beta\right)B^{o}}{\left(1-\beta\right)B^{o}}=\frac{\theta}{\left(1-\theta\right)}.$$

□

Next, we solve $\tilde{\mathcal{P}}\left(\tilde{B}\right)$.

Auxiliary Problem $\tilde{\mathcal{P}}\left(\tilde{B}\right)$: spot and chained consumption. Consider now the planner problem where at least some of both goods are consumed by the entrepreneur:

$$\tilde{\mathcal{P}}\left(\tilde{B}\right) = \max_{B\in\left[B^{\star}\left(\tilde{B}\right),\tilde{B}\right]} \tilde{\mathcal{P}}\left(B,\tilde{B}\right)$$

where

$$\tilde{\mathcal{P}}\left(B,\tilde{B}\right) \equiv (1-\theta)\log\left((1-\beta)B\right) + \theta\log\left(\mathcal{A}\left(\mu\left(B,\tilde{B}\right)\right)\cdot\mu\left(B,\tilde{B}\right) + S\left(B,\tilde{B}\right)\right),$$

subject to:

$$\mu\left(B,\tilde{B}\right) \equiv 1+\beta B-\tilde{B}$$

and

$$S\left(B,\tilde{B}\right) = \tilde{B}-B.$$

The derivative of the objective in $\tilde{\mathcal{P}}\left(\tilde{B}\right)$ is:

$$\tilde{\mathcal{P}}_B\left(B,\tilde{B}\right) = (1-\theta)\frac{1}{B} + \theta\frac{\mathcal{A}\left(\mu\left(B,\tilde{B}\right)\right)\left(1+\epsilon_{\mu}^{\mathcal{A}}\left(\mu\left(B,\tilde{B}\right)\right)\right)\beta-1}{\mathcal{A}\left(\mu\left(B,\tilde{B}\right)\right)\mu\left(B,\tilde{B}\right)+\tilde{B}-B}.$$

Recall that,

$$C\left(B,\tilde{B}\right) = \mathcal{A}\left(\mu\left(B,\tilde{B}\right)\right)\mu\left(B,\tilde{B}\right)+\tilde{B}-B.$$

and

$$Q\cdot C\left(B,\tilde{B}\right) = E\left(B,\tilde{B}\right) = 1-(1-\beta)B$$

Hence, using the definition of Q and q we rewrite:

$$\tilde{\mathcal{P}}_B\left(B,\tilde{B}\right) = (1-\theta)\frac{1}{B} - \theta Q\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}\left(\mu\left(B,\tilde{B}\right)\right)\right)\mathcal{A}\left(\mu\left(B,\tilde{B}\right)\right)}{1-(1-\beta)B}$$

We can multiply both sides by the ratio of $1-(1-\beta)B$ and divide by $(1-\beta)(1-\theta)$

and obtain:

$$\tilde{\mathcal{P}}_B\left(B,\tilde{B}\right)\frac{1-(1-\beta)B}{(1-\theta)(1-\beta)B}=\frac{1-(1-\beta)B}{(1-\beta)B}-\frac{\theta}{1-\theta}Q\left(\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\mathcal{A}(\mu)}{1-\beta}\right).$$

This function must have the same sign as $\tilde{\mathcal{P}}_B\left(B,\tilde{B}\right)$, since it was obtained by multiplication of positive numbers. The first term is decreasing in B. In turn, $Q_\mu \mu_B$ is increasing in B. Hence, as long as

$$\mathcal{A}(\mu)\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)=\mathcal{A}(\mu)+\mathcal{A}'(\mu)\mu$$

is decreasing in B, the second term is increasing. The second term is indeed decreasing in μ since its derivative is:

$$2\mathcal{A}'(\mu)+\mathcal{A}''(\mu)\mu<0,$$

where the sign follows immediately from the concavity and monotone decreasing properties of $\mathcal{A}$. Hence, the objective function $\tilde{\mathcal{P}}\left(B,\tilde{B}\right)$ is concave in B. Furthermore, since

$$Q\left(\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\mathcal{A}(\mu)}{1-\beta}\right)>1,$$

and we have an interior maximum in the region $B\in\left[B^{\star}\left(\tilde{B}\right),\tilde{B}\right]$, B must solve:

$$\frac{1-(1-\beta)B}{(1-\beta)B}=\frac{\theta}{1-\theta}Q\left(\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\mathcal{A}(\mu)}{1-\beta}\right),$$

and must be such that $B<B^{o}$.

Next, we establish properties regarding the limits of this function at $B^{\star}\left(\tilde{B}\right)$ and $\tilde{B}$. We start with $B^{\star}\left(\tilde{B}\right)$:

$$\lim_{B\downarrow B^{\star}\left(\tilde{B}\right)}\frac{1-(1-\beta)B}{(1-\beta)B}-\frac{\theta}{1-\theta}Q\left(\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\mathcal{A}(\mu)}{1-\beta}\right)=$$

$$\frac{1-(1-\beta)B^{\star}\left(\tilde{B}\right)}{(1-\beta)B^{\star}\left(\tilde{B}\right)}-\frac{\theta}{1-\theta}\left(\frac{1}{1-\beta}-\frac{\beta\left(1+\lim_{\mu\downarrow0}\epsilon_{\mu}^{\mathcal{A}}\right)\lim_{\mu\downarrow0}\mathcal{A}(\mu)}{1-\beta}\right),$$

where we used $\lim_{\mu\downarrow 0} Q = 1$. Appendix B, shows that $\lim_{\mu\downarrow 0} 1+\epsilon_{\mu}^{\mathcal{A}} = 0$ and that $\lim_{\mu\downarrow 0} \mathcal{A}(\mu) = \delta$. Thus, the limit of the objective function at the left boundary has the sign of:

$$\frac{1-(1-\beta)B^{\star}\left(\tilde{B}\right)}{(1-\beta)B^{\star}\left(\tilde{B}\right)} - \frac{\theta}{1-\theta}\frac{1-\beta\delta}{1-\beta} = \frac{1-(1-\beta)B^{\star}\left(\tilde{B}\right)}{(1-\beta)B^{\star}\left(\tilde{B}\right)} - \frac{1-(1-\beta)B^{o}}{(1-\beta)B^{o}}\frac{1-\beta\delta}{1-\beta}.$$

The function $B^{\star}\left(\tilde{B}\right) < B^{o}$ but $\frac{1-\beta\delta}{1-\beta} > 1$, hence the sign is ambiguous. The solution is at this corner if the function is negative.

Next, we consider the limit of the derivative of the objective at $\tilde{B}$. Evaluating the limits is immediate. Hence, if

$$\frac{1-(1-\beta)\tilde{B}}{(1-\beta)\tilde{B}} \geq \frac{\theta}{1-\theta}\left(\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}(\mu)\right)\mathcal{A}(\mu)}{1-\beta}\right)\Bigg|_{\mu=\left(1-(1-\beta)\tilde{B}\right)}$$

the solution is $B = \tilde{B}$. Moreover, we know that since

$$\left(\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)\mathcal{A}(\mu)}{1-\beta}\right) > 1,$$

the corner solution $\tilde{B}$ is chosen only if

$$\frac{1-(1-\beta)\tilde{B}}{(1-\beta)\tilde{B}} > \frac{1-(1-\beta)B^{o}}{(1-\beta)B^{o}}.$$

Collecting the results up to this point, we have the following Lemma.

Lemma 17. *The solution B_{ℓ}^{p} to $\tilde{\mathcal{P}}\left(\tilde{B}\right)$ is as follows.*

I. $B_{\ell}^{p} = \tilde{B}$ *if*

$$\frac{1-(1-\beta)\tilde{B}}{(1-\beta)\tilde{B}} \geq \frac{\theta}{1-\theta}\left(\frac{1-\beta\left(1+\epsilon_{\mu}^{\mathcal{A}}(\mu)\right)\mathcal{A}(\mu)}{1-\beta}\right)\Bigg|_{\mu=\left(1-(1-\beta)\tilde{B}\right)},$$

II. $B_{\ell}^{p} = B^{\star}\left(\tilde{B}\right)$ *if*

$$\frac{1-(1-\beta)B^{\star}\left(\tilde{B}\right)}{(1-\beta)B^{\star}\left(\tilde{B}\right)} \leq \frac{\theta}{1-\theta}\frac{1-\beta\delta}{1-\beta}.$$

III. *Otherwise, B_ℓ^p solves:*

$$\frac{1-(1-\beta)\,B_\ell^p}{(1-\beta)\,B_\ell^p} = \frac{\theta}{1-\theta} Q\left(B_\ell^p, \tilde{B}\right) \left(\frac{1-\beta\left(1+\epsilon_\mu^{\mathcal{A}}\left(\mu\left(B_\ell^p, \tilde{B}\right)\right)\right)\mathcal{A}\left(\mu\left(B_\ell^p, \tilde{B}\right)\right)}{1-\beta} \right)\Bigg|_{\mu=\left(1+\beta B_\ell^p - \tilde{B}\right)}.$$

The following Lemma characterizes threshold values for $\tilde{B}$ corresponding to the Lemma above:

Lemma 18. *The condition*

$$\frac{1-(1-\beta)\,B^\star\left(\tilde{B}\right)}{(1-\beta)\,B^\star\left(\tilde{B}\right)} \leq \frac{\theta}{1-\theta}\frac{1-\beta\delta}{1-\beta}$$

holds for all $\tilde{B} \geq \tilde{B}^i$ such that:

$$\frac{1-(1-\beta)\,B^\star\left(\tilde{B}^i\right)}{(1-\beta)\,B^\star\left(\tilde{B}^i\right)} = \frac{\theta}{1-\theta}\frac{1-\beta\delta}{1-\beta}.$$

The reverse inequality holds for $\tilde{B} \leq \tilde{B}^a$ such that:

$$\frac{1-(1-\beta)\,B^\star\left(\tilde{B}^a\right)}{(1-\beta)\,B^\star\left(\tilde{B}^a\right)} = \frac{\theta}{1-\theta}\left(\frac{1-\beta\left(1+\epsilon_\mu^{\mathcal{A}}(\mu)\right)\mathcal{A}(\mu)}{1-\beta}\right)\Bigg|_{\mu=\left(1-(1-\beta)\tilde{B}^a\right)}.$$

Proof. The proof follows immediately from the fact that:

$$\frac{1-(1-\beta)\,B^\star\left(\tilde{B}\right)}{(1-\beta)\,B^\star\left(\tilde{B}\right)},$$

is decreasing in $\tilde{B}$ and the function:

$$\frac{\theta}{1-\theta} Q\left(B, \tilde{B}\right) \left(\frac{1-\beta\left(1+\epsilon_\mu^{\mathcal{A}}\left(\mu\left(B, \tilde{B}\right)\right)\right)\mathcal{A}\left(\mu\left(B, \tilde{B}\right)\right)}{1-\beta} \right)\Bigg|_{\mu=(1-(1-\beta)B)}$$

increasing in B. □

Overall Solution. We showed above that:

$$\mathcal{P}^{\theta}\left(\tilde{B}\right)=\max\left\{\tilde{\mathcal{P}}\left(\tilde{B}\right),\underline{\mathcal{P}}^{\theta}\right\}.$$

The following Lemma shows that the solution B^p to the Planner's problem is never at $B^p=\tilde{B}$.

Lemma 19. *The planner never chooses $B^p=\tilde{B}$.*

Proof. To prove this Lemma observe that the left limit as $B\uparrow\tilde{B}$ satisfies

$$\frac{q-1-\epsilon^{\mathcal{A}}}{1-\beta}+1+\epsilon^{\mathcal{A}}>1+\epsilon^{\mathcal{A}},$$

where the inequality follows from $q>1$ and $\epsilon^{\mathcal{A}}<0$. As a consequence,

$$\frac{1-(1-\beta)\,\tilde{B}}{(1-\beta)\,\tilde{B}}\geq\frac{\theta}{1-\theta}\lim_{B\uparrow\tilde{B}}\Lambda\left(B,\tilde{B}\right)$$

implies

$$\frac{1-(1-\beta)\,\tilde{B}_t}{(1-\beta)\,\tilde{B}_t}>\frac{\theta}{1-\theta}\lim_{B\downarrow\tilde{B}_t}\Lambda\left(B,\tilde{B}\right).$$

Hence, although the derivative of the objective is discontinuous at $\tilde{B}$, we know that if the derivative is weakly increasing from the left and increasing from the right. This implies that $B=\tilde{B}$ is never an optimal choice. □

Next, observe $\tilde{\mathcal{P}}\left(\tilde{B}\right)$ has a compact-valued and continuous constraint correspondence with a continuous objective. Hence, it satisfies the conditions for the Theorem of the Maximum. In addition it is immediate to verify that:

$$\lim_{\tilde{B}\uparrow\frac{1-\theta\beta}{1-\beta}}\tilde{\mathcal{P}}\left(\tilde{B}\right)=\bar{\mathcal{P}}^{\theta}\text{ and }\lim_{\tilde{B}\uparrow\frac{1-\theta\beta}{1-\beta}}B^p(\tilde{B})=B^o.$$

We employ the Envelope Theorem on $\tilde{\mathcal{P}}\left(\tilde{B}\right)$. In the region where the solution to $\tilde{\mathcal{P}}\left(\tilde{B}\right)$ is not at a corner solution, the Envelope Theorem yields:

$$\tilde{\mathcal{P}}_{\tilde{B}}=\frac{\theta}{1-(1-\beta)\,B^p(\tilde{B})}Q\left(B^p(\tilde{B}),\tilde{B}\right)\left(1-\left(1+\epsilon^{\mathcal{A}}_{\mu}\left(\mu\left(B^p(\tilde{B}),\tilde{B}\right)\right)\right)\mathcal{A}\left(\mu\left(B^p(\tilde{B}),\tilde{B}\right)\right)\right)>0,$$

since $\tilde{B}$ appears directly through μ and S in the objective.

The function is strictly increasing in $\tilde{B}$ since $\epsilon_\mu^{\mathcal{A}} < 0$ and $\mathcal{A} < 1$. In the region where the function is at the lower corner of the constraint:

$$B = B^\star \left(\tilde{B}\right),$$

the value of this term is:

$$\tilde{\mathcal{P}}_{\tilde{B}} = \frac{\theta}{1-(1-\beta)\, B^\star \left(\tilde{B}\right)} (1-\delta) > 0$$

and

$$\tilde{\mathcal{P}}_B|_{B=B^\star(\tilde{B})} = \frac{1-\theta}{(1-\beta)\, B^\star \left(\tilde{B}\right)} - \frac{\theta\,(1-\beta\delta)}{1-(1-\beta)\, B^\star \left(\tilde{B}\right)}.$$

Thus, we have that the marginal objective value is:

$$\tilde{\mathcal{P}}_{\tilde{B}} + \tilde{\mathcal{P}}_B \left(B^p(\tilde{B})\right) B^\star_{\tilde{B}} \left(\tilde{B}\right) = \frac{1-\theta}{(1-\beta)\, B^\star \left(\tilde{B}\right)} \beta^{-1} + \frac{\theta\delta}{1-(1-\beta)\, B^\star \left(\tilde{B}\right)} > 0.$$

Hence, the envelope condition guarantees that $\tilde{\mathcal{P}}$ is increasing in $\tilde{B}$.

There exists a threshold $\tilde{B}^s$ such that:

$$\tilde{\mathcal{P}} \left(\tilde{B}^s\right) = \underline{\mathcal{P}}^\theta.$$

In particular, it solves:

$$\bar{\mathcal{P}}^\theta = \underline{\mathcal{P}}^\theta + \int_{\tilde{B}^s}^{\frac{1-\theta\beta}{1-\beta}} \left(\tilde{\mathcal{P}}_{\tilde{B}} + \tilde{\mathcal{P}}_B \left(B^p(\tilde{B})\right) B^\star_{\tilde{B}} \left(\tilde{B}\right)\right) d\tilde{B}.$$

Notice that $\underline{\mathcal{P}}^\theta$ is finite and lower than the value of the problem without constraints. $\tilde{\mathcal{P}}\,(0)$ tends to $-\infty$ and $\tilde{\mathcal{P}}\,(\overline{B})$ tends to the value of problem without constraints. This implies that $\tilde{\mathcal{P}}\,(0) < \underline{\mathcal{P}}^\theta < \tilde{\mathcal{P}}\,(\overline{B})$ and from continuity of the value function (Theorem of the Maximum) and the Intermediate Value Theorem, the existence of $\tilde{B}^s$ is guaranteed. The solution to $\tilde{B}^s$ follows from the fundamental theorem of calculus.

Recall that in the previous Lemma we had showed that $\tilde{B}$ is never a solution to $\tilde{\mathcal{P}}$ and the monotonicity implies the existence of the threshold $\tilde{B}^i$. These thresholds segment the intervals as given by the proposition. This concludes the proof of Proposition 8.

F.3 Discussion of Planner's Solution.

The solution to the Primal Problem reveals a novel insight. Namely, during transitions away from payments-chain crises, debt may be inefficiently high or inefficiently low. The intuition behind this is that, for a given debt level, the planner can increase output by reducing or increasing debt. By reducing debt, the planner distributes wealth toward the worker and frees credit lines inducing greater spot expenditures. Alternatively, by increasing debt, the planner distributes wealth toward the saver also stimulating spot expenditures. Both distributive policies distort equality, but increase efficiency by speeding up production.

The analytic expression for B^p reveals how the planner allocates expenditures, balancing productive efficiency against equality differently, depending on whether $\tilde{B}$ falls in several intervals. There are four intervals of values of $\tilde{B}$ where the planner's solution is qualitatively different. These actual critical values depend on the threshold points $\frac{1+\theta\beta}{1-\beta}$ and $\left\{\tilde{B}^s, \tilde{B}^i\right\}$ whose formulas are found in the proof.

The first interval of values is where $\tilde{B}$ is above a threshold level above which the SBL constraint is not binding. In particular, if $\tilde{B} \geq \frac{1+\theta\beta}{1-\beta}$, the planner sets $B = B_{ss}$, and production is efficient. Furthermore, the ratio of marginal utilities equals the ratio of Pareto weights—the efficient equality condition (14) is satisfied.

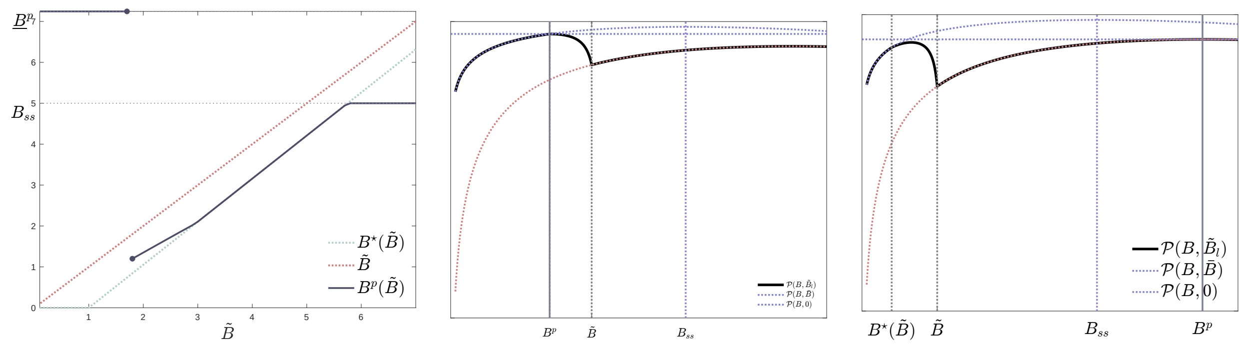
If $\tilde{B}$ is below the efficiency threshold, $\frac{1+\theta\beta}{1-\beta}$, the planner distorts either equality or productive efficiency. The novelty is that the planner has two ways to increase productive efficiency: the planner can increase efficiency on the margin, either by distributing wealth toward the worker if $B < \tilde{B}$ or by distributing wealth toward the saver if $B > \tilde{B}$. To see this, observe that in the region where $B < \tilde{B}$, any further reduction in debt translates, on the margin, into more spot expenditures by the worker and, thus, increases efficiency. In the region where $B > \tilde{B}$, any increase in debt translates, on the margin, into more spot expenditures by the saver, also increasing efficiency. Because of this ambivalent nature, the planner's objective function $\mathcal{P}\left(B, \tilde{B}\right)$ is not concave in B, leading to bang-bang solutions as $\tilde{B}$ varies.

Figure F.1 describes the economics of the Primal Problem, in the regions away from efficiency. The left panel plots B^p for different values of $\tilde{B}$. In the interval of values where $\tilde{B}$ falls in a second interval, i.e. when $\tilde{B} \in [\tilde{B}^i, \frac{1+\theta\beta}{1-\beta})$, the planner solution achieves productive efficiency, $\mathcal{A} = 1$, but the ratio of marginal utilities no longer equals the ratio of Pareto weights, as in (14). To understand why the planner does not distort production, observe that because $B^{\star}\left(\tilde{B}\right) < B_{ss}$, productive efficiency can only be achieved by setting $B^{p} = B^{\star}\left(\tilde{B}\right) < B_{ss}$. Setting debt at $B^{\star}\left(\tilde{B}\right)$, which is less than its steady-state value, implies that saver expenditures are less than in steady-state. Thus, in this interval the planner sacrifices equality by redistributing wealth toward the worker to guarantee productive efficiency. The planner's solution is at a corner because the derivative of his objective is discontinuous at $B^{\star}\left(\tilde{B}\right)$. This discontinuity follows from the payments-chain network structure: the property that $\mathcal{A}(0) < 1$, which captures that even when the chained expenditure ratio is zero, an individual chained order features a production delay.

When $\tilde{B}$ falls in the third interval, $\tilde{B} \in \left[\tilde{B}^{s}, \tilde{B}^{i}\right)$, the Planner begins to sacrifice productive efficiency. To maintain productive efficiency, for those levels of the SBL, the planner would have to redistribute even more wealth to the worker at the expense of saver expenditures. In that region, sacrificing equality alone is not worth it. Hence, the planner sets debt above its efficient level, thus $B^{\star}\left(\tilde{B}\right) < B^{p} < \tilde{B}$. Therefore, in this interval, the planner still redistributes wealth (relative to steady state) toward the agent facing the financial constraint, but the planner does partially sacrifice productive efficiency. The planner balances productive efficiency and equality, as given by equation (**??**).

When $\tilde{B}$ falls below an extreme value, in the fourth interval, $\tilde{B} < \tilde{B}^{s}$, the nature of the planner's solution changes dramatically. Increasing productive efficiency by redistributing wealth toward the worker requires the planner to set debt below $\tilde{B}$. However, when $\tilde{B} < \tilde{B}^{s}$, setting debt so low would require an extremely high sacrifice of saver consumption. Below that threshold, the planner prefers to redistribute wealth away from the worker, the constrained agent. Since the planner sets $B > \tilde{B}$, the planner induces only chained expenditures by the worker. Once the worker only makes chained expenditures, the SBL becomes irrelevant, so the planner chooses a constant debt level in this region. This debt level is higher than the unconstrained ideal debt level B_{ss} and given by Equation (**??**). Debt is higher than in steady state because the planner understands that a marginal increase in debt increases the wealth and, therefore, spot expenditures of savers. This increases productive efficiency.

The ambivalent nature of the planner's problem is central to understanding payments-chain crises. In typical models with pecuniary externalities, a planner wants to rebalance wealth toward the financially constrained agents to increase productive efficiency. Thus, typically, equality and productive efficiency reinforce each other. Here, for extremely low values of $\tilde{B}$, the planner switches to a policy mix where equality and efficiency are in conflict. The middle and right panels of Figure F.1 illustrate the change in the planner's strategy: In the middle panel, the planner prefers a value of debt where equality and efficiency complement. The right panel shows how he switches strategy as the SBL is tighter.



(a) Planner debt choice (b) $\mathcal{P}\left(\cdot,\tilde{B}\right)$ (moderate $\tilde{B}$) (c) $\mathcal{P}\left(\cdot,\tilde{B}\right)$ (tight $\tilde{B}$)

Figure F.1: Primal Planner Solution

Note: This figure reports a numerical example of the planner's solution. Figures are calculated using value function iteration: $\beta = 0.95$, $\delta = 0.9$, and $\theta = 0.75$. Panel (a) plots B^p as a function of values of $\tilde{B}$ in the range $[0.1, \theta + 0.1] * B^{star}$. Panels (b) and (c) plot the objective of the planner for different values of B in the range $[0, 1.8]$. Panel (b) fixes $\tilde{B}$ at 0.15 and Panel (c) at 0.08.

F.4 Descentralization with Taxes

To characterize the constrained efficient allocation, we demonstrate that Ramsey problem is equivalent to a problem where a planner chooses a path of debt respecting the payments technology and the desired expenditure rules of both agents.

For ease of exposition we reproduce both problems here. The Ramsey problem is:

Problem 10. (Ramsey Problem): *Given* B_0 *and* $\left\{\tilde{B}_t\right\}$*:*

$$\mathcal{V}^{Ramsey} = \max_{\left\{\tau^k_{t+1},\tau^c_t\right\}_{t\geq 0}} \sum_{t\geq 0} \beta^t \left[(1-\theta)\log\left(C_t\right) + \theta \log\left(S_t + X_t\right)\right],$$

subject to the creditor budget constraint and optimality:

$$\left(1+\tau^k_{t+1}\right) R^{-1}_{t+1} B_{t+1} + \left(1+\tau^c_t\right) C_t = B_t, \; \forall t \geq 0,$$

the entrepreneur budget constraint:

$$B_t + (1+\tau_t^c)\, E_t = R_{t+1}^{-1} B_{t+1} + Y,\ \forall t \geq 0,$$

the government's budget constraint:

$$\tau_{t+1}^k R_{t+1}^{-1} B_{t+1} + \tau_t^c\,(C_t + E_t) = 0,\ \forall t \geq 0,$$

and the three conditions of a competitive equilibrium: (i) Optimization, (ii) Asset and Goods market clearing, and (iii) the price consistency.

In turn, the Primal Problem is:

Problem 11. (Primal Planner Problem):*Taking* $\left\{\tilde{B}_t\right\}$ *as given:*

$$\mathcal{V}^{Primal} = \max_{\{B_t\}_{t\geq 0}} \sum_{t\geq 0} \beta^t \mathcal{P}\left(B_t, \tilde{B}_t\right)$$

where

$$\mathcal{P}\left(B, \tilde{B}\right) \equiv (1-\theta)\log\left((1-\beta)B\right) + \theta \log\left(S\left(B, \tilde{B}\right) + X\left(B, \tilde{B}\right)\right), \tag{86}$$

and taking as given $S\left(B, \tilde{B}\right)$ *and* $X\left(B, \tilde{B}\right)$ *as defined by Proposition 4.*

The Primal Problem differs from the competitive equilibrium in that the planner directly chooses the sequence of debt. The allocation and final goods is the same as the competitive equilibrium.

We will prove the following Proposition:

Proposition 9. (Primal and Ramsey Equivalence): *Let* $\left\{B_{t+1}^p\right\}_{t\geq 0}$ *be a solution to the Primal Planner Problem. The solution to the Ramsey problem achieves the same value as the Primal Planner Problem,* $\mathcal{V}^{Ramsey} = \mathcal{V}^{Primal}$*. The solution to the Primal Planner Problem can be implemented setting* $(1+\tau_0^c) = B_0/B_0^p$ *and a sequence of debt taxes given in the proof below.*

The proposition asserts that the solution to the Ramsey Problem can be obtained from the solution of a Primal Problem where the planner directly chooses the sequence of debt.[60]

[60] This relation follows because the constraint set in the Primal Problem includes the constraints of the

If a solution to the primal can be implemented with taxes, it solves the Ramsey Problem. The proposition shows that this is the case. The implementation of the Primal Problem is possible because the Ramsey planner can set capital taxes to induce a desired path of debt $\{B_t\}$. Controlling B_t is key to the implementation.[61]

Before we proceed to the proof, we establish the following useful intermediate result:

Lemma 20 (Creditor's Solution with Taxes)**.** *Let* $\{\tau^k_{t+1}, \tau^c_t, \tau^\ell_{t+1}\}_{t\geq 0}$ *be a sequence of taxes in the Ramsey Problem. The solution to the creditor's problem is:*

$$C_t = (1-\beta)B^o_t, \quad (1+\tau^c_t)C_t = (1-\beta)B^o_t$$

where $B^o_t = B_t/(1+\tau^c_t)$ *and* $B^o_{t+1} = \hat{R}_{t+1}\beta B^o_t$ *with:*

$$\hat{R}_{t+1} \equiv \frac{R_{t+1}}{(1+\tau^k_{t+1})} \cdot \frac{1}{(1+\tau^c_{t+1})/(1+\tau^c_t)}$$

Proof. The creditor's problem with taxes is:

$$V_t = \sum_{s=0}^{\infty} \beta^s \log(C_{t+s})$$

subject to $(1+\tau^k_{t+1})R^{-1}_{t+1}B_{t+1} + (1+\tau^c_t)C_t = B_t$. Dividing both sides by $(1+\tau^c_t)$ and defining $B^o_t \equiv B_t/(1+\tau^c_t)$, we can re-write the budget constraint as:

$$\hat{R}^{-1}_{t+1}B^o_{t+1} + C_t = B^o_t \tag{87}$$

where $\hat{R}_{t+1} \equiv \frac{R_{t+1}}{(1+\tau^k_{t+1})} \cdot \frac{1}{(1+\tau^c_{t+1})/(1+\tau^c_t)}$.

With log utility and a budget constraint like (87), the creditor's optimal expenditure rule is:

$$C_t = (1-\beta)B^o_t$$

Ramsey Problem. This is immediate since market clearing in the asset market and the budget balance, implies, by Walras's law, that the resource constraint holds. In the implementation, capital taxes and a period-zero expenditure tax are required by the Ramsey planner to distort the evolution of debt and replicate the solution to the Primal Problem.

[61] Time separability also implies that expenditure taxes are redundant after $t = 0$. All that the Ramsey planner needs is a sequence of capital taxes to control the path of debt and the labor tax to redistribute the capital tax receipts toward the worker. The planner only needs expenditure taxes at $t = 0$ because debt is predetermined at $t = 0$.

which yields $B^o_{t+1} = \hat{R}_{t+1}\beta B^o_t$. Since $B_t = (1+\tau^c_t)B^o_t$, we have $C_t = (1+\tau^c_t)(1-\beta)B^o_t$ as claimed. □

F.5 Proof Ramsey-Primal Equivalence - Proposition 9

Proof. To characterize the constrained efficient allocation, the strategy is to show that any solution to the Ramsey Problem satisfies the constraints of the Primal Problem (Step 1). Then, we show that any solution to the Primal Problem can be induced by a proper tax sequence $\{\tau^k_{t+1}, \tau^c_t, \tau^\ell_{t+1}\}_{t\geq 0}$ (Step 2).

Step 1: From Ramsey to Primal - Deriving the Resource Constraint. Adding the two agents' budget constraints we arrive at:

$$(1+\tau^c_t)(S_t + q_t X_t + C_t) + \tau^k_{t+1}\frac{B_{t+1}}{R_{t+1}} = 1 - \tau^\ell_{t+1}.$$

Subtracting the government budget constraint we arrive at:

$$(1+\tau^c_t)(S_t + q_t X_t + C_t) + \tau^k_{t+1}\frac{B_{t+1}}{R_{t+1}} - \left(\tau^k_{t+1}\frac{B_{t+1}}{R_{t+1}} + \tau^c_t(C_t + S_t + q_t X_t) + \tau^\ell_{t+1}\right) = 1 - \tau^\ell_{t+1}$$

simplifying to:

$$S_t + q_t X_t + C_t = 1.$$

From Lemma 20, $C_t = (1-\beta)B^o_t$, given $B^o_t = B_t/(1+\tau^c_t)$. Therefore:

$$S_t + q_t X_t = 1 - (1-\beta)B^o_t. \tag{88}$$

Thus, setting $E_t = S_t + q_t X_t$ we have that the entrepreneurs expenditures are:

$$E_t = 1 - (1-\beta)B^o_t. \tag{89}$$

The optimality of spot and chained expenditures is still given by the conditions in Proposition 4.

Using $q_t = \mathcal{A}(\mu_t)^{-1}$ and $\mu_t = q_t X_t = \mathcal{A}(\mu_t)^{-1}X_t$. These are exactly the constraints of the Primal Problem with B^o_t as the state variable instead of B^p_t.

Step 2: Tax Implementation of the Ramsey problem. Let $\{B^p_t\}_{t\geq 0}$ solve the Primal

Problem. We now show how to construct taxes that implement this as a Ramsey equilibrium. Set the initial consumption tax:

$$(1+\tau_0^c) = B_0/B_0^p$$

This ensures $B_0^o = B_0/(1+\tau_0^c) = B_0^p$.

For the Primal solution to be implementable, we need both agents' Euler equations to hold. The creditor's Euler equation:

$$\frac{(1-\beta)B_{t+1}^o/(1-\beta)B_t^o}{\beta} \cdot \frac{(1+\tau_{t+1}^c)}{(1+\tau_t^c)} \cdot (1+\tau_{t+1}^k) = R_{t+1} \tag{90}$$

The entrepreneur's Euler equation (using the notation from Proposition 2):

$$\frac{(S_{t+1}+X_{t+1})/(S_t+X_t)}{\beta} \cdot \frac{(1+\tau_{t+1}^c)}{(1+\tau_t^c)} \cdot \frac{\tilde{q}_{t+1}^B(B_{t+1}^o)}{\tilde{q}_t^E(B_{t+1}^o)} = R_{t+1} \tag{91}$$

where the marginal price ratio captures the payment chain friction.

Eliminating $[(1+\tau_{t+1}^c)/(1+\tau_t^c)]\beta R_{t+1}$ from both equations:

$$(1+\tau_{t+1}^k) = \frac{B_t^o}{B_{t+1}^o} \cdot \frac{S_{t+1}+X_{t+1}}{S_t+X_t} \cdot \frac{\tilde{q}_{t+1}^B(B_{t+1}^o)}{\tilde{q}_t^E(B_{t+1}^o)}$$

Since $S_t + X_t = 1-(1-\beta)B_t^o$ and using average prices $Q_t = (S_t + q_t X_t)/(S_t + X_t)$:

$$\tau^k(B_t^o, B_{t+1}^o, \tilde{B}_t, \tilde{B}_{t+1}) = \frac{B_t^o}{B_{t+1}^o} \cdot \frac{1-(1-\beta)B_{t+1}^o}{1-(1-\beta)B_t^o} \cdot \frac{Q(B_t^o, \tilde{B}_t)}{Q(B_{t+1}^o, \tilde{B}_{t+1})} \cdot \frac{\tilde{q}_{t+1}^B(B_{t+1}^o)}{\tilde{q}_t^E(B_{t+1}^o)} - 1 \tag{92}$$

The equilibrium interest rate follows from (90):

$$R_{t+1} = \frac{(1+\tau_{t+1}^c)}{(1+\tau_t^c)} \cdot \frac{B_{t+1}^o}{\beta B_t^o} \cdot (1+\tau^k(B_t^o, B_{t+1}^o, \tilde{B}_t, \tilde{B}_{t+1}))$$

Step 3: Indeterminacy of Consumption Taxes Note that consumption taxes $\{\tau_t^c\}_{t\geq 1}$ are indeterminate. To see this, substitute the creditor's Euler equation into their budget constraint:

$$\frac{(1-\beta)B_t^o}{(1-\beta)B_{t+1}^o}\beta\left[\frac{(1+\tau_t^c)}{(1+\tau_{t+1}^c)}\right]B_{t+1} + (1+\tau_t^c)(1-\beta)B_t^o = B_t$$

Since $B_t = (1+\tau_t^c)B_t^o$, this is satisfied for any sequence of consumption taxes. The government budget constraint then determines:

$$\tau_{t+1}^{\ell} = -\tau_{t+1}^{k}\frac{B_{t+1}}{R_{t+1}} - \tau_t^c\left((1-\beta)B_t^o + \frac{1-(1-\beta)B_t^o}{Q(B_t^o, \tilde{B}_t)}\right)$$

Setting $\tau_t^c = 0$ for all $t \geq 1$ yields the simple implementation:

$$\tau_{t+1}^{\ell} = -\frac{\tau^k(B_t^o, B_{t+1}^o, \tilde{B}_t, \tilde{B}_{t+1})}{1+\tau^k(B_t^o, B_{t+1}^o, \tilde{B}_t, \tilde{B}_{t+1})} \cdot \beta B_t^o$$

This completes the proof that $\mathcal{V}^{Ramsey} = \mathcal{V}^{Primal}$. □

F.6 Derivation of the Bocola Effect - Proof of Proposition 5

Let's define the indirect social utility function of the Planner Problem with expenditures $\mathcal{P}^g$:

$$\mathcal{P}^g\left(B,\tilde{B},G^s,G^x\right) = (1-\theta)\log\left(C^s\left(B\right)\right)+\theta\log\left(S^w\left(B,\tilde{B},G^s,G^x\right)+X^w\left(B,\tilde{B},G^s,G^x\right)\right).$$

The following functions determine the allocation:

$$\begin{aligned}
C^s\left(B\right) &\equiv (1-\beta)\,B\\
E^w\left(B,G^s,G^x\right) &\equiv 1-E^s\left(B\right)-G^x-G^s\\
S^w\left(B,\tilde{B},G^s,G^x\right) &\equiv \min\left\{\max\left\{\tilde{B}-B,0\right\},E^w\right\}\\
X^w\left(B,\tilde{B},G^s,G^x\right) &\equiv \frac{E^w-\min\left\{\max\left\{\tilde{B}-B,0\right\},E^w\right\}}{q}\\
q\left(B,\tilde{B},G^s,G^x\right) &\equiv \mathcal{A}^{-1}\left(\mu\right)\\
\mu\left(B,\tilde{B},G^s,G^x\right) &\equiv G^x+X^w q.
\end{aligned}$$

The worker's total consumption is:

$$C^w = S^w\left(B,\tilde{B},G^s,G^x\right)+X^w\left(B,\tilde{B},G^s,G^x\right).$$

Because

$$\mathcal{Y} = C^w + C^S + G^s + \frac{G^x}{q\left(\mu\right)},$$

but $C^s\left(B\right)$ is independent of the government's expenditure, we obtain: $d\mathcal{Y} = dC^w + dG^s + d\left(G^x/q\left(\mu\right)\right)$.

Thus, we have that the government's expenditure multiplier for expenditure of type $i = x, s$ relates to the worker's consumption as follow:

$$\mathcal{M}^s\left(B,\tilde{B}\right) \equiv \frac{d\mathcal{Y}}{dG^s} = \frac{dC^w}{dG^s}+1$$

and

$$\begin{aligned}\mathcal{M}^x\left(B,\tilde{B}\right) &\equiv \frac{d\mathcal{Y}}{dG^x}=\frac{dC^w}{dG^x}+\frac{1}{q}-\frac{G^x}{q^2}\frac{dq}{d\mu}\frac{d\mu}{dG^x}\\ &= \frac{dC^w}{dG^x}+\mathcal{A}\left(\mu\right).\end{aligned}$$

where the last equation is because we are interested in obtaining the government's infinitesimal multiplier, evaluated at $G^x = G^s = 0$. We have defined these multipliers as:

$$\mathcal{M}^i\left(B,\tilde{B}\right)\equiv \left.\frac{d\mathcal{Y}}{dG^i}\right|_{G^x=G^s=0},$$

and relating to the indirect social utility function, the change in the objective is:

$$\frac{\theta}{C^w}\frac{dC^w}{dG^i},$$

for expenditure i.

Case 1. All spot consumption $B < B^{\star}$. If there is only spot consumption:

$$\begin{aligned}C^s\left(B\right) &\equiv \left(1-\beta\right)B\\ C^w\left(B,\tilde{B},G^s,G^x\right) &\equiv 1-\left(1-\beta\right)B-G^x-G^s\\ X^w\left(B,\tilde{B},G^s,G^x\right) &\equiv 0\\ q\left(B,\tilde{B},G^s,G^x\right) &\equiv \mathcal{A}^{-1}\left(\mu\right)\\ \mu\left(B,\tilde{B},G^s,G^x\right) &\equiv G^x.\end{aligned}$$

Since worker consumption is independent of q, we have that:

$$\frac{dC^w}{dG^i}=-1$$

for both $i\in\{x,s\}$. In turn, we have that

$$\mathcal{M}^s\left(B,\tilde{B}\right)=\frac{dC^w}{dG^s}+1=0.$$

Likewise, for chained expenditures we have:

$$\mathcal{M}^{x}\left(B,\tilde{B}\right)=\ \frac{dC^{w}}{dG^{x}}+\mathcal{A}\left(\mu\right)\left(1+\epsilon_{\mu}^{\mathcal{A}}\right)=-\left(1-\mathcal{A}\left(\mu\right)\right)+\epsilon_{\mu}^{\mathcal{A}}<0.$$

Case 2. Some chained consumption. If there are some chained expenditures:

$$\begin{aligned}
C^{s}\left(B\right)&\equiv\ \left(1-\beta\right)B\\
E^{w}\left(B,G^{s},G^{x}\right)&\equiv 1-\left(1-\beta\right)B-G^{x}-G^{s}\\
C^{w}\left(B,\tilde{B},G^{s},G^{x}\right)&\equiv S^{w}+X^{w}\\
S^{w}\left(B,\tilde{B},G^{s},G^{x}\right)&\equiv\ \max\left\{\tilde{B}-B,0\right\}\\
X^{w}\left(B,\tilde{B},G^{s},G^{x}\right)&\equiv\frac{E^{w}-\max\left\{\tilde{B}-B,0\right\}}{q}\\
q\left(B,\tilde{B},G^{s},G^{x}\right)&\equiv\mathcal{A}^{-1}\left(\mu\right)\\
\mu\left(B,\tilde{B},G^{s},G^{x}\right)&\equiv G^{x}+X^{w}q.
\end{aligned}$$

Rewriting the last three identities using $\mathcal{A}\left(\mu\right)$ instead of q we have

$$\begin{aligned}
X^{w}&\equiv\mathcal{A}\left(\mu\right)\left(E^{w}-\max\left\{\tilde{B}-B,0\right\}\right)\\
\mu&\equiv G^{x}+E^{w}-\max\left\{\tilde{B}-B,0\right\}.
\end{aligned}$$

Substituting E^{w} we have, naturally,

$$\mu=1-\underbrace{\left(\left(1-\beta\right)B+G^{s}+\max\left\{\tilde{B}-B,0\right\}\right)}_{\text{spot exp.}}. \tag{93}$$

From here we obtain that:

$$\frac{dX^{w}}{dG^{x}}=\ -\mathcal{A}\left(\mu\right)+\mathcal{A}'\left(\mu\right)\frac{d\mu}{dG^{x}}\left(E^{w}-\max\left\{\tilde{B}-B,0\right\}\right).$$

Since $\frac{d\mu}{dG^{x}}=\frac{dS^{w}}{dG^{x}}=0$, we have that:

$$\frac{dC^{w}}{dG^{x}}=-\mathcal{A}\left(\mu\right).$$

Hence, the government multiplier for chained expenditures is:

$$\mathcal{M}^x\left(B,\tilde{B}\right) = \epsilon_\mu^{\mathcal{A}} < 1.$$

Next, observe that:

$$X^w \equiv \mathcal{A}\left(\mu\right)\left(1-\left(1-\beta\right)B - G^x - G^s - \max\left\{\tilde{B}-B,0\right\}\right).$$

Hence,

$$\begin{aligned} dX^w &= -\mathcal{A}\left(\mu\right)dG^s + \mathcal{A}'\left(\mu\right)\left(\frac{X^w}{\mathcal{A}\left(\mu\right)}\right)d\mu \\ &= -\mathcal{A}\left(\mu\right)dG^s + \mathcal{A}\left(\mu\right)\epsilon_\mu^{\mathcal{A}}\left(\frac{X^w}{\mathcal{A}\left(\mu\right)}/\mu\right)d\mu. \end{aligned}$$

The second line follows from:

$$\mathcal{A}\left(\mu\right)\mu \equiv G^x + X^w.$$

Also following this condition, we have that:

$$\mathcal{A}\left(\mu\right)\left(1+\epsilon_\mu^{\mathcal{A}}\right)d\mu \equiv dX^w.$$

Combining the differentials evaluated at $G^x = 0$, we obtain:

$$d\mu = -dG^s.$$

Hence,

$$\frac{dC^w}{dG^s} = \frac{dX^w}{dG^s} = -\mathcal{A}\left(\mu\right)\left(1+\epsilon_\mu^{\mathcal{A}}\right).$$

Following the relationship with the fiscal multiplier, we obtain:

$$\mathcal{M}^s\left(B,\tilde{B}\right) = 1 - \mathcal{A}\left(\mu\right) - \mathcal{A}\left(\mu\right)\epsilon_\mu^{\mathcal{A}}.$$

Summary. We summarize the results:

$$\frac{dC^w}{dG^x} = \begin{cases} -1 & B < B^\star\left(\tilde{B}\right) \\ \\ -\mathcal{A}\left(\mu\right) & B > B^\star\left(\tilde{B}\right) \end{cases}$$

and

$$\frac{dC^w}{dG^s} = \begin{cases} -1 & B < B^\star\left(\tilde{B}\right) \\ \\ -\mathcal{A}\left(\mu\right)\left(1+\epsilon_\mu^{\mathcal{A}}\right) & B > B^\star\left(\tilde{B}\right). \end{cases}$$

Finally, recall that the government multipliers relate to the change in consumption as follows:

$$\mathcal{M}^s\left(B,\tilde{B}\right) \equiv \frac{d\mathcal{Y}}{dG^s} = \frac{dC^w}{dG^s} + 1$$

and

$$\mathcal{M}^s\left(B,\tilde{B}\right) \equiv \frac{dC^w}{dG^x} + \mathcal{A}\left(\mu\right)$$

with:

$$\frac{d\mu}{dG^x} = \begin{cases} 1 & B < B^\star\left(\tilde{B}\right) \\ \\ 0 & B > B^\star\left(\tilde{B}\right). \end{cases}$$

Therefore, adding terms:

$$\mathcal{M}^x\left(B,\tilde{B}\right) = \begin{cases} -\left(1-\mathcal{A}\left(\mu\right)\left(1+\epsilon_\mu^{\mathcal{A}}\right)\right) & B < B^\star\left(\tilde{B}\right) \\ \\ 0 & B > B^\star\left(\tilde{B}\right). \end{cases}$$

Hence,

$$\mathcal{M}^{s}\left(B,\tilde{B}\right)=\begin{cases}0 & B<B^{\star}\left(\tilde{B}\right)\\ \\ 1-\mathcal{A}\left(\mu\right)\left(1+\epsilon_{\mu}^{\mathcal{A}}\right) & B>B^{\star}\left(\tilde{B}\right).\end{cases}$$