

NBER WORKING PAPER SERIES

DATA-DRIVEN AUTOMATION

Maryam Farboodi
Andrew J. Koh
Anchi Xia

Working Paper 35320
<http://www.nber.org/papers/w35320>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2026

We thank Daron Acemoglu, David Autor, Tom Cunningham, Eric Gao, Basil Halperin, Fredric Kong, Anton Korinek, Stephen Morris, Arjun Ramani, Pascual Restrepo, Phil Trammell, and Iván Werning for helpful conversations. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Maryam Farboodi, Andrew J. Koh, and Anchi Xia. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Data-Driven Automation

Maryam Farboodi, Andrew J. Koh, and Anchi Xia

NBER Working Paper No. 35320

June 2026

JEL No. J31, O33, O4

ABSTRACT

We build a dynamic model of data-driven automation in which data (i) is heterogeneous and task-specific; (ii) accumulates endogenously as a byproduct of economic activity; and (iii) exhibits spillovers such that data generated by one task can augment the productivity of another. Along the transition path of automation, data plays a dual role in simultaneously augmenting the productivity of already-automated tasks and expanding the automation frontier. We derive tight conditions for the economy to be partially versus fully automated in the long-run. In the latter case, automation exhibits rich short-run dynamics that depend on the pattern of data spillovers but is always slow in the long-run: the share of tasks produced by labor decays asymptotically as a power law in time. We show that the economy is generically inefficient and analyze how a planner optimally tilts the direction of data accumulation. With endogenous capital accumulation, data-driven automation generates explosive growth but stagnant long-run wages.

Maryam Farboodi
Massachusetts Institute of Technology
MIT Sloan School of Management
and NBER
m.farboodi@gmail.com

Anchi Xia
Massachusetts Institute of Technology
bryantxia2435@gmail.com

Andrew J. Koh
Massachusetts Institute of Technology
Department of Economics
AJKoh96@gmail.com

1 Introduction

Rapid advancements in artificial intelligence have sparked intense debate about the future of work and the trajectory of automation across the economy. For instance, Dario Amodei, CEO of Anthropic has warned of a ‘white collar bloodbath’ (Axios, 2025). A recent US Senate report has likewise warned that nearly 100 million US jobs might disappear over the next decade (U.S. Senate Committee, 2025). What are we to make of such claims? While previous waves of technological change primarily displaced routine tasks (Autor et al., 2003), modern AI systems, trained on vast corpora of human media, increasingly demonstrate capabilities that rival human performance across a growing range of cognitively demanding tasks. For instance, the ability of AI systems to complete long software tasks—those that would take t units of time for human software engineers—has been growing exponentially since the release of GPT-2 in 2019 (Kwa et al., 2025). Such technologies raise fundamental questions about which tasks of the economy will be automated, how quickly this transition will occur, whether automation will be limited to certain domains or become pervasive across all aspects of economic life, and what this might mean for growth and wages in the short and long run.

Central to understanding the process of automation is recognizing that modern machine learning systems are fundamentally data-driven: AI systems require extensive training data to perform tasks, and their performance continues to improve even in the ‘big data’ regime of tens of trillions of tokens. Put differently, AI and data are strong complements. Moreover, there is a growing trend of finetuning models on task-specific data.¹ All of this data does not arise in a vacuum, but was rather generated as a byproduct of economic activity, giving rise to the following *data-automation feedback loop*:

Data $\rightarrow$ *Automation*. Modern AI systems rely on staggering amounts of data. Pre-trained models ingest virtually the entire corpus of human work, and are increasingly fine-tuned on task-specific data—workflows, norms, and job-specific details e.g., call transcripts, a law firm’s legal cases, or a hedge fund’s investment process—to improve their performance in specific domains. A striking empirical regularity—so-called ‘neural scaling laws’—is that more abundant training data and/or model parameters leads to predictable decreases in the loss across a wide suite of model architectures.² This shapes *task-specific* capital productivity and, in turn, drives the path of equilibrium automation.

¹Examples include Capoot (2025) for legal tasks, Jiang et al. (2023) for medicine, and (Wu et al. (2023) and ‘BloombergGPT’ for financial analysis.

²Some examples of losses are coding mistakes, wrong medical diagnoses, malfunctioning of robots or autonomous vehicles, or wrong mathematical proofs.

Automation $\rightarrow$ *Data*. Where does this data come from? Up to now, large-language models have been largely pre-trained on the internet. However, a growing view within the computer science community is that in order for AI models to perform economically valuable tasks presently done by humans, simply training larger models with more compute may not be enough: they need to be trained on *task-specific data*³ that might encode tacit know-how and contextual knowledge that can only be gained from learning-by-doing. For instance, this is the explicit ambition of firms that build self-driving vehicles, and have become increasingly prevalent—OpenAI reportedly expects the “entire economy” to become a “Reinforcement Learning Machine,” where AI systems might train on recordings of how professionals in all fields handle day-to-day work on their devices (The Information, 2025). Indeed, there are increasing attempts to collect data in the workplace (Paul and Horwitz, 2026) and Silicon Valley is beginning to build virtual work environments — a ‘virtual machine outfitted with an email inbox, a Slack account, some coding tools and a web browser’ (NYT, 2025) — in which firms might train an AI system to automate the *entirety* of a white collar job. Hence, the intensity and composition of production—and automation—across the economy will shape task production and, in so doing, drive differential data accumulation.

Motivated by these developments, we build a simple macroeconomic model to incorporate this data-automation feedback loop. Our model takes three distinctive features of data seriously. First, data is a byproduct of economic activity—different kinds of data are differentially valuable across tasks. Second, data is a long-lived asset that accumulates endogenously as a by product of economics activity. Third, data exhibits spillovers—data generated by one task can augment productivity of another task.

We characterize the scale and speed of automation and capital concentration in equilibrium, as a function of the structural parameters of the economy. Our model delivers the following key insights:

1. *Long-run automation*: The degree of automation in the long-run depends jointly on cross-task substitutability and the extent of diminishing returns to data.
2. *Wages*: In equilibrium labor ‘trains its own replacement’. Long-run wages stagnate whenever the economy is fully-automated in the limit e.g., whenever tasks are gross complements.
3. *Speed of automation*: The path of equilibrium automation can be rapid or imbalanced, and depends richly on the network of cross-task spillovers. But automation eventually slows down: sans capital accumulation, the proportion of tasks produced by labor decays according to a power law in time.

³For an articulation of the view that better task-specific data is needed for automation, see <https://www.mechanize.work/blog/sweatshop-data-is-over/>

4. *Contagion*: Automation can spread contagiously across the economy in the presence of cross-task data spillovers driven by either transfer learning or overlaps in primitive ‘skills’.
5. *Core-periphery automation*: The network structure of data spillovers can generate imbalanced automation in the short-to-medium run in which a small number of full-automated sectors (‘core’) are overproduced and generate data that feed-back into those sectors while bottlenecks to spillovers slow automation in the wider economy (‘periphery’).
6. *(In)efficiency*: The composition of data accumulation is generically inefficient, and the direction of inefficiency i.e., whether data-rich sectors are accumulating data too slowly or quickly depends on the degree of cross-task substitution.
7. *Explosive growth*: Data-automation feedback loops drive long-run explosive growth in the presence of endogenous capital accumulation.

We now explain these results in a more detail. We first shut down capital accumulation and analyze an economy in data autarky where data is purely task-specific and there are no cross-task spillovers. We describe conditions under which the economy achieves full versus partial automation in the long-run, and study the transition path of automation. We show that the direction and limit of automation are determined by two key parameters: σ , which governs the substitutability across tasks, and η , which governs the returns to scale of data on capital productivity.

When $\sigma \leq 1/\eta$, the degree to which tasks are complementary exceeds the speed at which capital productivity runs into diminishing returns from additional data. We show that the economy is then fully automated in the long run (**Proposition 1**). The economy asymptotes towards a *balanced data path* while task-specific productivities continuously improve. Intuitively, the capital allocation is not too skewed towards data-rich tasks and hence the ratio of task capital productivities approaches a constant. In such economies, data-driven automation ensures that the set of tasks produced by labor decline quickly enough so that our economy escapes the cost-disease trap (Baumol, 1967).

Alternatively, when $\sigma > 1/\eta$, the degree of task complementarity is dominated by the speed at which capital productivity runs into diminishing returns. In this case, we show that the economy can remain perpetually on an *imbalanced data path*—equilibrium capital allocation tilts toward data-rich tasks which, in turn, makes them more data-rich. In such cases, the ratio of data in task i (data-rich) on task j (data-poor) approaches infinity so that automation can, in fact, fall along the equilibrium path and the economy is only partially automated in the limit, with capital concentrated in a small set of highly productive tasks (**Proposition 2**).

What does this mean for wages? We show that in the case $\sigma < 1/\eta$ regime, wages initially rise but eventually stagnate as the productivity gains from data-augmented capital is offset by endogenous displacement from a growing automation boundary (**Proposition 3**). By contrast, when tasks are sufficiently substitutable and automation is partial ($\sigma > 1/\eta$), long-run wages grow unboundedly. Together, this implies, somewhat surprisingly, that cross-task complementarities hurt long-run wages. Indeed, the ‘partial equilibrium’ logic is that for a *fixed* vector of capital productivities, decreasing σ (thus making tasks more complementary) increases wages. But when data is dynamically accumulated in equilibrium, an increase in cross-task complementarity means that production of ‘data-poor’ tasks—those that are yet to be automated—are higher in equilibrium, and this, in turn, displaces labor and drives wages down until the ‘productivity effect’ (from automated sectors being more productive) and the ‘displacement effect’ (from labor performing a shrinking share of tasks) exactly offset each other.

Furthermore, we develop tight non-asymptotic bounds on the speed of data-driven automation (**Proposition 4**). We show that although equilibrium automation can be potentially rapid in the short term, it takes a long time for full equilibrium automation i.e., for (almost) all tasks to be performed by capital in equilibrium. In particular, we show—sans capital accumulation—the share of tasks produced by labor decays as a power law in time. Thus, although data-driven automation is powerful in the limit, on its own it is also *slow*: there remains a ‘fat tail’ of tasks that continue to be performed by labor—not because capital is incapable of doing so, but because, given the prevailing price of compute and the limited efficiency of capital at producing those tasks, it is more worthwhile to produce them with labor.

We then generalize the model by introducing cross-task data spillovers into our economy. We say that the economy is *connected* if, for any two tasks i, j , there exists a ‘positive measure path’ of tasks through which data can spillover from i to j . Connectedness reflects the idea that even though tasks i and j may not directly benefit from each other’s data, we can find a chain of intermediate tasks $i := i_1 \rightarrow i_2 \rightarrow i_3 \rightarrow \dots \rightarrow j$ that are beneficiaries of the preceding task’s data. This series of spillovers ensure a kind of contagion through quantities produced in general equilibrium—an increase in $i = i_1$ ’s data augments i_2 ’s factor productivity that raises its equilibrium quantity. This increases the speed at which data on task i_2 accumulates, which then augments i_3 ’s factor productivity, and so on.⁴ We show that connectedness is sufficient to guarantee full automation in the long-run independent of the elasticity of substitution σ , return-to-scale η , and the initial condition on the data stock across tasks (**Proposition 5**). These, ‘local data spillovers’ can, in effect, drive automation by substituting for global cross-task complementarities. We further show that connectedness is suffi-

⁴We model data spillovers via a graphon (a dense continuous graph) so connectedness comprises of a ‘positive measure path’.

cient to guarantee a balanced data path and a balanced composition of tasks produced in the long run.

The network structure of cross-task spillovers also matter enormously for automation dynamics in the short-run. We illustrate this by specializing our model to *core-periphery automation* in which a subset of densely connected sectors (‘core’) are data rich, and as they are produced in equilibrium, data generated feeds back into the core, further augmenting their capital productivity. By contrast, there are only weak spillovers to remaining sectors that comprise the broader economy (‘periphery’). Although our previous long-run results state that this is sufficient to guarantee full long-run automation, how long might this take? We show that patterns of spillovers can generate imbalanced short-run automation in which a subset of tasks/sectors are rapidly automated but the wider economy is still dominated by labor. We analytically characterize the time at which the periphery sectors begin to be automated: it hinges on both the (i) strength of within-core linkages that reduce equilibrium incentives to automate peripheral sectors; and (ii) the severity of the bottlenecks in how quickly data flows from the core to periphery.

Next, we address the efficiency of data-driven automation and show that equilibrium automation is generically inefficient (**Proposition 6**). We show that along the transition path, the planner emphasizes the equilibrium direction of data convergence or divergence. When tasks are sufficiently complementary ($\sigma \leq 1/\eta$), the planner wishes to initially tilt the allocation of capital further *toward* data-poor tasks relative to the equilibrium allocation—this speeds up data accumulation for bottlenecked tasks at the cost of changing the equilibrium allocation of capital. Conversely, when tasks are strong substitutes ($\sigma > 1/\eta$) the planner wishes to tilt the allocation of capital further *away from* data-poor tasks relative to the equilibrium allocation. This offers a new perspective on how market forces can distort the *direction* of data production because firms do not internalize the impact of their task production on the evolution of the aggregate data stock.

Finally, we introduce endogenous capital accumulation into our model. We show that even in the case in which tasks are strongly complementary, additional data delivers steeply diminishing returns, and capital depreciates quickly, the fact that data (i) augments productivity of already-automated tasks; (ii) expands the automation boundary;⁵ and (iii) does not depreciate imply a growth explosion in finite time (**Proposition 7**). In contrast to the ‘supply-side singularity’ result of Nordhaus (2021), labor and capital are not necessarily gross substitutes at any point in time, but automation will eventually give rise to a representation of the economy where the effective substitutability is governed by the perfect substitutability between these two factors at the task level.

⁵See, e.g. Solow (1957); Acemoglu and Restrepo (2018); Aghion et al. (2017).

Connection to the literature. Our work relates closely to the task model (Zeira, 1998) and recent work analyzing the trajectory as well as impact of automation (Acemoglu and Restrepo, 2018; Aghion et al., 2017; Korinek and Suh, 2024; Trammell, 2025; Demirer et al., 2025). Different from these papers, our focus is on how task-specific data accumulation—determined both in general equilibrium, as well as via local spillovers—drives the *speed*, *direction*, and *limit* of automation across the economy. More related is Jones and Liu (2024) who show that expanding automation and growing capital productivity can jointly produce balanced growth if intermediate goods are gross complements. Jones and Liu (2024) embody technological change via an R&D process in the tradition of Romer (1986)—by contrast, our model takes seriously the idea that data rather than labor (in the form of human scientists) is the key input—and bottleneck—to modern AI systems, and it is the ensuing path of equilibrium data accumulation that delivers rising capital productivity.⁶

Data-automation feedback loops are qualitatively distinct from previous analyses of big data as being used to predict demand or improve product quality (Jones and Tonetti, 2020; Farboodi and Veldkamp, 2021).⁷ This is because the data-automation feedback loop emerges when data is specifically used for *training* AI models that can be duplicated at scale both within and across firms, and does not depreciate over time.⁸ We think this qualitatively distinguishes data-automation feedback loops from previous analyses e.g., predicting demand (Bajari et al., 2019).⁹

Finally, our analysis of local spillovers across nearby products, produced by tasks that share a subset of their tasks, is informed by modern machine learning systems that exhibit a degree of local ‘transfer learning’.¹⁰ To capture this, we endow our task space with a measure of distance, and tasks that are ‘nearby’ have more similarities in their production routine and thus exhibit more spillovers. Our formalization employs the language of graphons developed in mathematics (Borgs et al., 2008), which has recently been drawn on to study network games (Parise and Ozdaglar, 2023) and contagion (Erol, Parise, and Teytelboym, 2023; Koh and Morris, 2022). Our setting is substantially different because the intensity of spillover is jointly determined by ‘local’ and ‘global’ general equilibrium forces, and—as is typical in macroeconomics—

⁶Our focus is on the data dimension of modern AI systems and, in so doing, abstract unemployment; see Costinot and Werning (2023); Guerreiro et al. (2022); Beraja and Zorzi (2025) for analyses of policy remedies.

⁷These models build upon the pioneering work of Wilson (1975).

⁸See, Chatterji, Rock, and Talamas (2025); Hadfield and Koh (2025) for possibilities for how AI agents might be incorporated within firms.

⁹Of course, on a high level all data is used for prediction; indeed, even with large-language models or reinforcement learning algorithms, training data serves to augment next-token prediction. However, we think this is economically distinct from previous analyses of big data.

¹⁰See, e.g., Dwivedi and Roig (2019) who develop a measure of task similarity by assessing the performance of a model trained on different tasks. A recent interview by employees of Anthropic echo this ambition for large-language models (Patel, 2025b).

in the long-run data for all tasks diverge to infinity. Thus, it is their *relative speeds* that determines equilibrium automation.

2 A Data-Augmented Task Model

We begin by developing a simple model of production with task-specific data accumulation.

Final good and intermediate tasks. There is a task space $X = [0, 1]$. A final good is competitively produced by aggregating intermediate tasks according to the usual CES technology

$$Y = \left(\int_{i \in X} y_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad (1)$$

where y_i is the quantity produced of task $i \in X$. $\sigma > 1$ corresponds to gross substitutes, $\sigma < 1$ corresponds to gross complements, and the limit of $\sigma = 1$ corresponds to the neutral, Cobb-Douglas case.

Factors of production. We consider two factors of production, capital and labor. In our baseline model, capital and labor are fixed at $K > 0$ and $L > 0$ respectively.¹¹ We emphasize that the ‘labor block’ in our model is deliberately stylized such as to focus on the impact of how data accumulation might augment capital productivity heterogeneously across tasks; see, e.g., [Acemoglu, Kong, and Restrepo \(2025\)](#); [Autor and Thompson \(2025\)](#) for analyses of how automation can have differential occupational impacts.

Data. Data is a state variable that governs how productive capital is across tasks. Data is task-specific: the data that is used to train an AI model to build financial models is different from that used to train a surgery robot. We use $\mathbf{D}_t \in \mathcal{D}$ to denote the (infinite-dimensional) vector of time- t data across tasks, and write D_{it} to denote the cumulative data at time t from producing task i .¹²

Tasks. The production function of task i is linear in capital and labor, and is given by:

$$y_i = \psi^L l + \psi_i^K(\mathcal{A}_i(\mathbf{D})) \cdot k$$

¹¹In [Section 7](#) we present an extension with capital accumulation, where we introduce a standard household block that governs the consumption-savings decision in our economy. There we analyze how endogenous capital accumulation interacts with data-driven automation.

¹² $\mathcal{D}$ is the space of (infinite-dimensional) data stock that is the space of measurable functions from $X \rightarrow \mathbb{R}_+$ and for $\mathbf{D} \in \mathcal{D}$, $i \in X$ we write $D_i := \mathbf{D}(i)$ to denote the stock of data for task i .

where $\mathcal{D} \mapsto \mathcal{A}_i(\mathbf{D}) \in \mathbb{R}_+$ is a *data aggregator* tracking how the full vector of data translates into the *effective data* available for task i . We expand on properties of $\mathcal{A}_i(\mathbf{D})$ in detail in Section 2.1.

As capital and labor are perfect substitutes, each task will be performed using the input that is more productive for that specific task. When a task is produced with capital in equilibrium, we say that it is *automated*. We emphasize that automation is an equilibrium notion which depends jointly on technology (ψ) and factor prices (r and w). To keep track of the degree of automation in our economy, it is useful to introduce the following equilibrium object.

Definition 1 (Automation boundary). The time- t *automation boundary* $\gamma_t \in \text{Int}(X)$ denotes the threshold task that is automated—tasks $i \leq \gamma_t$ are produced via capital, and tasks $j > \gamma_t$ are produced via labor.¹³

Finally, to incorporate diminishing return to data, we impose the following assumptions on how data shape the marginal productivity of capital.

Assumption 1 (Productivity of data within and across tasks). The vector of *effective data* determines the marginal productivity of capital in each task $i \in X$ according to the following functional form,

$$\psi_i^K \left(\underbrace{\mathcal{A}_i(\mathbf{D})}_{\text{effective data}} \right) = f_i \cdot (\mathcal{A}_i(\mathbf{D}))^\eta$$

where $\eta \in (0, 1)$, common across all tasks, captures the degree of diminishing returns of data in capital productivity and $f_i > 0$ is a task-specific scalar coefficient that weakly decreases in the task index i . We use $\mathbf{f}$ to denote the vector of task-specific scalars, and assume that it is bounded above and below.

An alternate interpretation of f_i is that it encodes *differential speeds* of data accumulation per unit of task produced. We discuss this interpretation momentarily along with introducing spillover of data across tasks to capture the degree of transfer learning as documented in computer science.

2.1 Evolution of Data and Cross-Task Spillovers

The initial stock of data is given by $\mathbf{D}_0$, which is bounded, continuous, and strictly decreasing in task index. The stock of data evolves according to the following law of

¹³We break ties in favor of automation without loss of generality.

motion:

$$D_{it} = D_{i0} + \int_0^t y_{is} ds. \quad (2)$$

Thus, the accumulation of data generated by task i is simply the production of i . This reflects the key channel through which economic forces that determine equilibrium task quantities feed back into data accumulation. Indeed, in addition to text and image inputs (e.g., a legal firm’s past cases), modern AI training procedures are increasingly adept at taking in video inputs of the physical tasks (e.g., surgical tasks (Kim et al., 2024)) being performed and learning how to imitate accordingly. Note here that, different from Farboodi and Veldkamp (2021); Jones and Tonetti (2020), data does not depreciate—if an AI system augments the productivity of a given task, it can always do so.¹⁴

Note that Equation (2) specifies the *task-specific data accumulation*, i.e., how D_t is determined. Next, we explain the determination of $\mathcal{A}_i(\mathbf{D})$, i.e., how data is employed in production. Here, we assume that although data accumulation is task-specific, the data accumulated for a given task might be valuable for other tasks— $\mathcal{A}_i(\mathbf{D})$ can depend on D_j for $j \neq i$. That is, our economy might exhibit *cross-task spillovers*.

To describe patterns of spillovers across tasks, we define the directed graphon as the measurable function¹⁵ $W : [0, 1]^2 \rightarrow \mathbb{R}_+$. The data aggregator for task i under W is:

$$\mathcal{A}_i(\mathbf{D}) = \int_{j \in X} W(i, j) \cdot D_j dj.$$

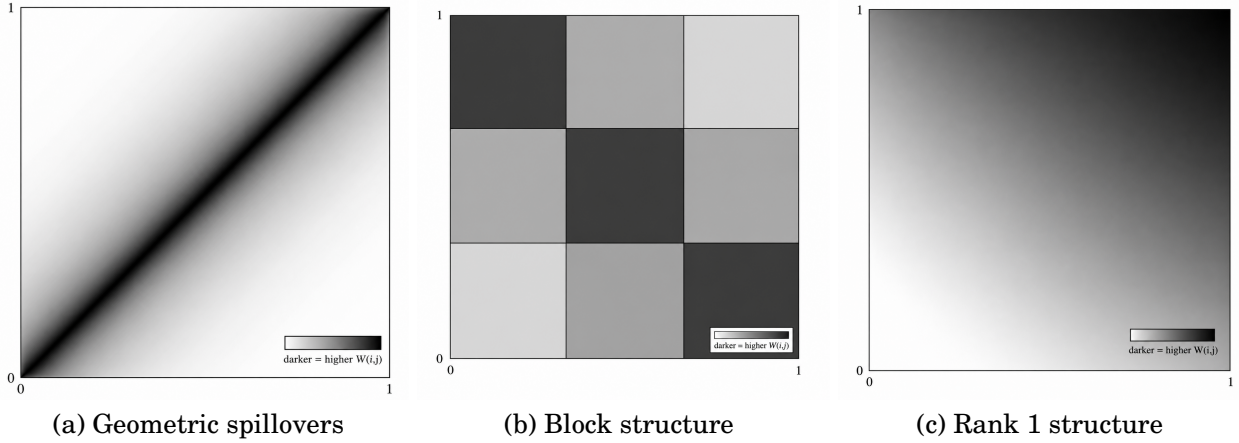
$W(i, j)$ denotes how much the productivity of task i is influenced by the data of task j and the ‘effective data’ that feeds into task i ’s capital productivity simply integrates over tasks $j \in X$. There are two ways to interpret our graphon W . First, it might capture the idea that two tasks $i, j \in X$ might share underlying skills such that data on how task i is performed might improve the performance of AI systems on those underlying skill that, in turn, improves the productivity of task j . Second, the graphon W might simply capture the degree of ‘transfer learning’—the ability of AI systems to generalize to new domains. This remains very much an empirical question fiercely debated among computer scientists.¹⁶ Our model of graphons keeps the shape and

¹⁴For instance, the depreciation of data in Farboodi and Veldkamp (2021) arises from a changing state so past data is less predictive of future states. However, we view a given task $i \in X$ as fixed; we discuss how our model might accommodate changing composition of tasks across the economy a la Acemoglu and Restrepo (2018); Korinek and Suh (2024) in Section 7.

¹⁵Graphons are sometimes defined with the condition of symmetry, but things are well-defined without it.

¹⁶For example, Demis Hassabis the CEO of DeepMind recently said in an interview that: “For example, when these large models improve at coding, that can actually improve their general reasoning.

Figure 1: Examples of cross-task spillovers



magnitude of data spillovers flexible, allowing us to analyze implications for the direction and limit of automation without taking a particular stance on the exact form of transfer learning.

Figure 1 illustrates three examples of graphons:

- (a) *Geometric spillovers*: endow the task space X with the Euclidean distance and data spillovers are *local* in the sense that task i benefits more from data generated by the production of nearby tasks than those further away.¹⁷ This specification is natural when the index i orders tasks by technological or occupational similarity. Data generated in one task is then most useful for nearby tasks whose inputs, workflows, or prediction problems are similar.
- (b) *Block structure*: tasks are partitioned into subsets (‘blocks’). Link strength is constant across each pair of blocks, so that tasks in the same origin block have identical spillovers to tasks in the same destination block. This nests finite-sector models as a special case. This specification is natural when tasks can be grouped into sectors or jobs that are broadly similar. For example, data may travel easily within a sector, while cross-sector spillovers depend mainly on the pair of sectors involved.
- (c) *Rank-one structure*: some tasks are systematically stronger sources of outflows, while others are systematically stronger than inflows. Then our graphon is $W(i, j) = u(i)v(j)$. For instance, ‘mathematical reasoning’ might be upstream

There is evidence of some transfer, although we would like a lot more evidence of that.” (Patel, 2025b).

¹⁷See Koh and Morris (2022), who also endow the graphon with geometry to study the speed of game-theoretic contagion, and more broadly the literature in probability theory on random geometric graphs.

of many other tasks (stronger outflows, high ν) while ‘writing’ might be receptive to a broad source of data (stronger inflows, high u).

We now have all the necessary elements to formally define the equilibrium in this economy, provided in Definition 2.

Definition 2 (Equilibrium). At each time t , given the data stock D_t , an equilibrium comprises of a rental rate r_t , a wage w_t , and allocations of capital and labor across tasks such that (i) the final good producer is maximizing her time- t profit; and (ii) intermediate task producers are maximizing their time- t profits.

2.2 Discussion

Before presenting our results, it will be helpful to discuss our main assumptions governing data-driven automation.

D1: Data as a source of heterogeneous capital productivity. The performance of modern AI systems can, to a first order, be broadly attributed to (i) advances in compute; (ii) algorithmic improvements; and (iii) increases in training data (Ho et al., 2024). While cheapening compute and algorithmic improvements impact capital productivity in a homogeneous manner (via the price of capital r), differences in the availability of task-specific data might explain differences in performance of AI systems across economic tasks.

Indeed, an interpretation of $f = (f_i)_{i \in X}$ (recall f_i is the constant governing how well data maps to capital productivity at task i) is that it encodes differential rates of data accumulation.¹⁸ There are several reasons why tasks might differ in their rates of data accumulation. For instance, tasks for which f_i is low might correspond to those that can only be done by AI systems trained on long multi-step procedures with sparse reward signals (e.g., low ‘bits per sample’) (Ord, 2025; Patel, 2025a). Tasks for which f_i is high might correspond to those that can be completed by AI models trained on denser e.g., video data (Wiedemer et al., 2025). Likewise, some tasks might be more easily verifiable than others, thereby corresponding to a quicker rate of data accumulation per unit produced (Karpathy, 2025). Our model offers a language to formalize this ‘generation-verification gap’ in which equilibrium forces determine the cost of generating data for different tasks, as well as the quantities actually generated.

D2: Productivity improvements via data. We make three remarks about Assumption 1, illustrated in Figure 2. First, our functional form is motivated by the

¹⁸That is, our economy is *identical* to another economy in which the functional form of capital productivity is the same across tasks, but different tasks accumulate data at different rates per unit of output.

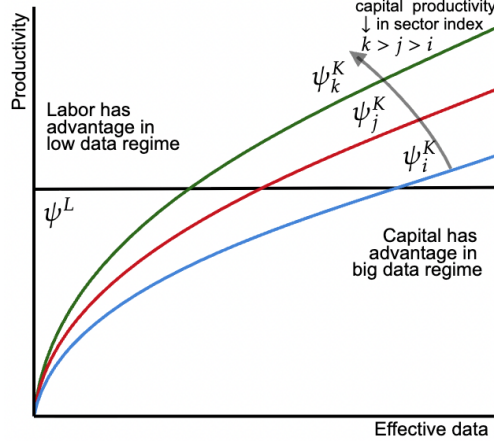


Figure 2: Illustration of ψ_i^K vs ψ^L

literature in computer science (see, e.g., [Hoffmann et al. \(2022\)](#) and a large body of subsequent work) demonstrating that the loss—a measure of imperfection for a task—diminishes polynomially with the amount of data used for training large language models. Heuristically, they find that the loss approximately takes the functional form: $\text{LOSS} \simeq \frac{1}{N^\alpha} + \frac{1}{D^\beta}$ for $\alpha, \beta > 0$ where D is data and N is the number of parameters which is typically linear in the amount of floating point operations (‘compute’).¹⁹ A growing view from computer science is that, given compute has been growing far more quickly than data (on which models can be trained),²⁰ the performance of AI systems—especially on specific economic tasks—might be increasingly bottlenecked by the scarcity of high quality data ([Kim et al., 2025](#); [Narayanan and Kapoor, 2025](#)). Moreover, building competent AI systems for the real economy (as opposed to, say, benchmarks) might not only require large and powerful pretrained models, but also data on the specific economic context within which such systems are deployed. Our model takes this seriously, and approximates the ‘data-constrained’ regime in which N is large so $\text{LOSS} \approx 1/D^\beta$. Under the (*substantial!*) assumption that economic productivity of a given task is the reciprocal of the statistical loss (perhaps raised to some power) we exactly recover the assumed functional form.²¹

Second, the economic content of [Assumption 1](#) reflects an extreme version of the ideas that (i) humans are more *sample efficient* than modern AI systems—in the ‘low data’ regime where training data for tasks is low, labor is more productive than capital; and that (ii) AI systems *saturate more slowly*—they are able to exploit vast amounts of data to drive loss down so that in the ‘big data’ regime, the opposite holds.

¹⁹[Hoffmann et al. \(2022\)](#) estimate this for the Huber loss; see also [Besiroglu et al. \(2024\)](#).

²⁰For instance, the training compute of frontier models have been growing at 4–5 times per year ([Sevilla and Roldán, 2024](#)).

²¹See also [Farboodi and Veldkamp \(2021\)](#) where the quality of data is a strictly decreasing function of the quadratic loss from prediction.

Third, we have chosen to let capital productivity grow unboundedly with data because what ultimately matters for equilibrium automation is the task cost r/ψ^K , which would also go to 0 with capital accumulation (as in [Acemoglu and Restrepo \(2018\)](#); [Aghion et al. \(2017\)](#)) holding capital productivity fixed. The key difference is that our model with heterogeneous and endogenous data accumulation allows us to speak to the *ratio* of cross-task productivity, i.e., ψ_i^K/ψ_j^K which drives the speed and limit of automation.

D3: Raw versus labeled data Our model is one of endogenous data accumulation as a byproduct of economic activity. This is already the case for certain industries such as self-driving cars in which data accumulates with production. We think our model of endogenous data accumulation well-approximates such tasks. For other tasks, however, current AI systems are not yet able to perform them—no matter how much computational power is thrown at the problem. For such tasks, the AI industry has resorted to hiring humans to produce data on which such systems can be trained *separately* rather than as a byproduct. For instance a wave of new tech firms are hiring lawyers, doctors, bankers, and journalists to lend their expertise in labeling datasets AI systems can be trained on ([Wall Street Journal, 2025](#)). We think this might simply be a ‘cold start’ problem that will be eventually overcome with enough human labeling and/or as workplace surveillance technologies improve and AI systems become increasingly able to learn from multimodal domains e.g., video (see [Cullen, Li, and Li \(2025\)](#)). But as soon as the task can be performed with capital, we expect equilibrium quantities to drive how task data is accumulated in equilibrium.²²

D4: Cross-task data spillovers. A striking example of cross-task spillover is the impressive performance of pre-trained models on multitask language understanding. A single model is capable of answering questions from computer security to business ethics ([Hendrycks et al., 2020](#)). However, we might expect a large-language model to be fine-tuned on data on how to compose emails to also excel at copywriting, but not at mathematical proofs or steering an autonomous vehicle. This is consistent with recent evidence that shows models finetuned on a classification task exhibits positive transfer within the same ‘kind’ of task (but in out-of-domain data), with more mixed performance across different kinds of tasks ([Yang et al., 2024](#)). By varying the network of transfer learning W , we will be able to link the speed, direction, and limit of automation under different degrees of transfer learning.

D5: Data-driven technological change. A distinctive feature of our model is the vector of task-specific data determine equilibrium allocation. It will, however, be

²²Our model can accommodate this ‘cold start’ in which, in the absence of enough data capital productivity is zero by writing $\psi_i^K(\mathcal{A}) = \max\{0, f_i \cdot \mathcal{A}^\eta - C_i\}$ where $C_i > 0$ is some task-specific constant such that capital requires $[C_i/f_i]^{1/\eta}$ units of effective data to be able to produce it.

helpful to connect our model to more ‘standard’ models of technological change in the Hicksian tradition (Hicks, 1932) formalized by Acemoglu (2002).

Define the following *capital productivity index*:

$$\Psi_K(\gamma_t, \boldsymbol{\psi}_t^K) := \left(\frac{1}{\gamma_t} \int_0^{\gamma_t} [\psi_{jt}^K]^{\sigma-1} dj \right)^{\frac{1}{\sigma-1}} \quad (3)$$

which is a function of the automation boundary γ_t and the profile of task productivities $\boldsymbol{\psi}_t^K := (\psi_{it}^K)_i$. By substituting our expressions for capital and labor productions into (1), output can be written in terms of our capital productivity index as follows:

$$Y_t = \left(\gamma_t^{\frac{1}{\sigma}} (\Psi_K(\gamma_t, \boldsymbol{\psi}_t) \cdot K)^{\frac{\sigma-1}{\sigma}} + (1 - \gamma_t)^{\frac{1}{\sigma}} (\psi_L \cdot L)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \quad (4)$$

Labor and capital augmenting technological change correspond to improvements in ψ_L and Ψ_K , holding the automation boundary γ_t fixed. A key feature of our model is that data accumulation simultaneously (i) increases capital productivities so that *fixing* the automation boundary γ_t , it is capital-augmenting; and (ii) *endogenously changes* the automation boundary γ_t which must equilibrate labor and capital markets. This connects our model to the important work of Acemoglu and Restrepo (2018); Aghion et al. (2017) who model automation only on the extensive margin (capital productivity is either 0 or 1), and with an exogenous movement of the capital boundary given by technology. Our model is also distinct from a recent paper of Jones and Liu (2024) who also allow both ‘intensive’ and ‘extensive’ margins for automation via an R&D process à la Romer (1986).²³ Differently, our model takes seriously recent trends in AI progress that (i) automation is at least, in part, *data-driven*; (ii) data is heterogeneous and *task-specific*, and (iii) data is *produced endogenously* as a by-product of production.

3 Automation in Data Autarky

We start by considering the case where there are no data spillovers across tasks i.e., $\mathcal{A}_i(\mathbf{D}) = D_i$ for all $i \in X$.²⁴ Lemma 1 characterizes the static equilibrium—the au-

²³We might view a special case of our model with uniform spillovers i.e., for each stock of data $\mathbf{D}$, $\mathcal{A}_i(\mathbf{D}) = \mathcal{A}_j(\mathbf{D})$ for all $i, j \in X$ as recovering a version of ‘learning by doing’ à la Arrow (1962); Romer (1986) in which productivity is a one-dimensional function of past production. By taking heterogeneous and task-specific data seriously, we will move between the ‘data autarky’ (Section 3) with no spillovers, and economies with richer patterns of spillovers (Section 4).

²⁴Although the data-autarky case is technically not nested within our graphon framework, it can be viewed as the limit of local averaging kernels: for $\epsilon > 0$, define $W^\epsilon(i, j) := \frac{\mathbf{1}_{\{|i-j| \leq \epsilon\}}}{\mu_{\{k \in [0,1]: |i-k| \leq \epsilon\}}}$. Then for

tomation boundary, factor prices, and allocations—and establishes that the equilibrium path is regular in the sense that at each time $t \geq 0$ task capital productivities are indeed decreasing in task index.

Lemma 1 (Static equilibrium characterization and regular paths). Suppose that there are no data spillovers. If capital productivity is weakly decreasing in task index, then:

- (i) **Automation boundary.** There exists a unique task $\gamma_t \in (0, 1)$ satisfying Equation (5) such that tasks in $[0, \gamma_t]$ produce with capital and tasks in $(\gamma_t, 1]$ produce with labor.

$$\psi_{\gamma_t}^K = \underbrace{\psi^L \left(\frac{\int_0^{\gamma_t} (\psi_j^K)^{\sigma-1} dj}{(1-\gamma_t)(\psi^L)^{\sigma-1} K} \right)^{\frac{1}{\sigma}}}_{r/w} \quad (5)$$

- (ii) **Factor prices and allocations.** Capital and labor allocations are given by Equations (6) and (7):

$$K_i = \frac{(\psi_i^K)^{\sigma-1}}{\int_0^{\gamma_t} (\psi_j^K)^{\sigma-1} dj} K \quad \text{for all } i \in [0, \gamma_t] \quad (6)$$

$$L_i = \frac{L}{1-\gamma_t} \quad \text{for all } i \in (\gamma_t, 1] \quad (7)$$

Rental rate of capital and wages are given by Equations (8) and (9):

$$r = Y^{\frac{1}{\sigma}} \left(\frac{\int_0^{\gamma_t} (\psi_j^K)^{\sigma-1} dj}{K} \right)^{\frac{1}{\sigma}} \quad (8)$$

$$w = Y^{\frac{1}{\sigma}} \left(\frac{(1-\gamma_t)(\psi^L)^{\sigma-1}}{L} \right)^{\frac{1}{\sigma}}. \quad (9)$$

If, in addition, the initial data stock $\mathbf{D}_0$ is continuous and strictly decreasing in task index, then there exists an equilibrium path such that for all times t , capital productivity is weakly decreasing in task index.

Note that these equilibrium allocations and factor price characterizations apply in more general settings when monotonicity of capital productivity across tasks may not

every $D \in L^1([0, 1])$, $\mathcal{A}_i^e(\mathbf{D}) \rightarrow D_i$ as $\epsilon \downarrow 0$ for a.e. $i \in [0, 1]$. Thus $\mathcal{A}_i(\mathbf{D}) = D_i$ is the local-kernel limit of the aggregators $\mathcal{A}_i^e$.

hold. In such cases, simply replace the integrating region with the set of tasks that are cheaper to produce by capital. We next characterize how the extent of automation in this economy changes over time.

Lemma 2 (Direction of the automation boundary). The automation boundary, γ_t , is weakly increasing in time if and only if the following inequality holds

$$(\sigma - 1) \int_0^{\gamma_t} \left(\frac{f(j)}{f(\gamma_t)} \right)^{1/\eta} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{2\sigma-1-1/\eta} dj \leq \sigma \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \quad (10)$$

Notice that Equation (10) is more slack for small σ . In particular, it holds trivially when $\sigma \leq 1$ as the left hand side is weakly negative while the right hand side is positive. In other words, if tasks are gross complements, the automation boundary moves to the right monotonically and more tasks are automated over time.

To gain some preliminary intuition for this result, start from a given automation boundary $\hat{\gamma}$. Let time progress but hold fixed the automation boundary. Data accumulation acts as a purely capital augmenting technological change, so capital becomes relatively cheaper compared to labor due to standard Baumol forces when tasks are gross complements (Baumol, 1967; Nordhaus, 2008). Consequently, capital productivity of the cutoff task must decrease so as to equilibrate (recall $r/w = \psi_{\gamma}^K/\psi^L$). On the other hand, we know data accumulation improves capital productivity of all tasks over time, so the only way for this relation to hold is for γ_t to increase, as higher indexed tasks are less productive for capital.

In Subsections 3.1 and 3.2 we leverage Lemma 2 to precisely characterize the evolution of the automation boundary, before discussing the implications of data-driven automation for equilibrium wages in Subsection 3.3.

3.1 Low Substitutability Across Tasks

Let us first analyze the case where the elasticity of substitution across tasks σ is smaller than the reciprocal of the degree of diminishing returns to data, $1/\eta$. Note that $1/\eta > 1$ so this nests both the case where tasks are gross complements and the case where tasks are gross substitutes but not too similar.

Our first result is that in this case, full automation is inevitable—for any initial condition on data D_0 and any heterogeneity in how tasks might differentially benefit from data f , so long as Assumption 1 holds, the economy is always automated in the limit.

Proposition 1 (Full automation and balanced data). If $\sigma \leq 1/\eta$:

(i) **Full automation:** The economy is fully automated in the limit, that is, $\lim_{t \rightarrow \infty} \gamma_t = 1$. If tasks are additionally gross complements (i.e., $\sigma \leq 1$), then automation increases monotonically.

(ii) **Balanced data:** For any $i, j \in \text{Int}(X)$ where $i < j$:

$$\lim_{t \rightarrow \infty} \frac{D_{it}}{D_{jt}} = \left(\frac{f_i}{f_j} \right)^{\frac{\sigma}{1-\sigma\eta}}.$$

The intuition behind **Proposition 1** is quite different for the gross complements case ($\sigma \leq 1$) and for the gross substitutes case with ($1 \leq \sigma \leq 1/\eta$). We discuss each in turn.

First consider the case when tasks are gross complements ($\sigma \leq 1$). Towards a contradiction, suppose automation is upper-bounded by some level $\bar{\gamma} \in \text{Int}(X)$. This implies data accumulation must become purely capital augmenting after some large time. By the usual Baumol logic, the capital share of the economy as well as the relative price of capital to labor will both collapse to zero. On the other hand, we know data for all tasks grows unboundedly so capital becomes infinitely productive which yields the contradiction

$$\underbrace{\psi_{\bar{\gamma}}^K}_{\rightarrow \infty} = \underbrace{(r_t/w_t)}_{\substack{\rightarrow 0 \\ \text{(Baumol)}}} \cdot \psi^L$$

since the per unit cost for producing any task $i \in \text{Int}(X)$ cannot equilibrate between capital and labor. Indeed, the logic of this argument is quite general and applies more broadly:

Remark 1. *When tasks are gross complements ($\sigma \leq 1$) our economy is fully automated in the limit for any functional form for capital productivity $\psi_i^K : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ that grows unboundedly as (effective) data grows unboundedly.*

In addition to the economy being fully automated in the limit, the capital share will also converges to 1. This is the result of a tug of war between intensive capital productivity improvements and the extensive expansion of the automation boundary. Whereas the former drives the usual Baumol cost disease, shrinking the size of the economy attributable to capital, the latter directly enlarges the capital share which is most readily seen from **Equation (4)**. The previous argument implies that automation dominates the Baumol effect in the limit; **Figure 3a** illustrates numerically that this is also the case along the transition. Intuitively, automation occurs quickly enough because capital productivity growth in the marginal task is keeping pace with growth among inframarginal tasks. At each time t , capital allocation is *biased towards data-poor tasks* (**Figure 4a**), partially offsetting inequities in task production due to current levels of capital productivity and resulting in a balanced data path in the long run.

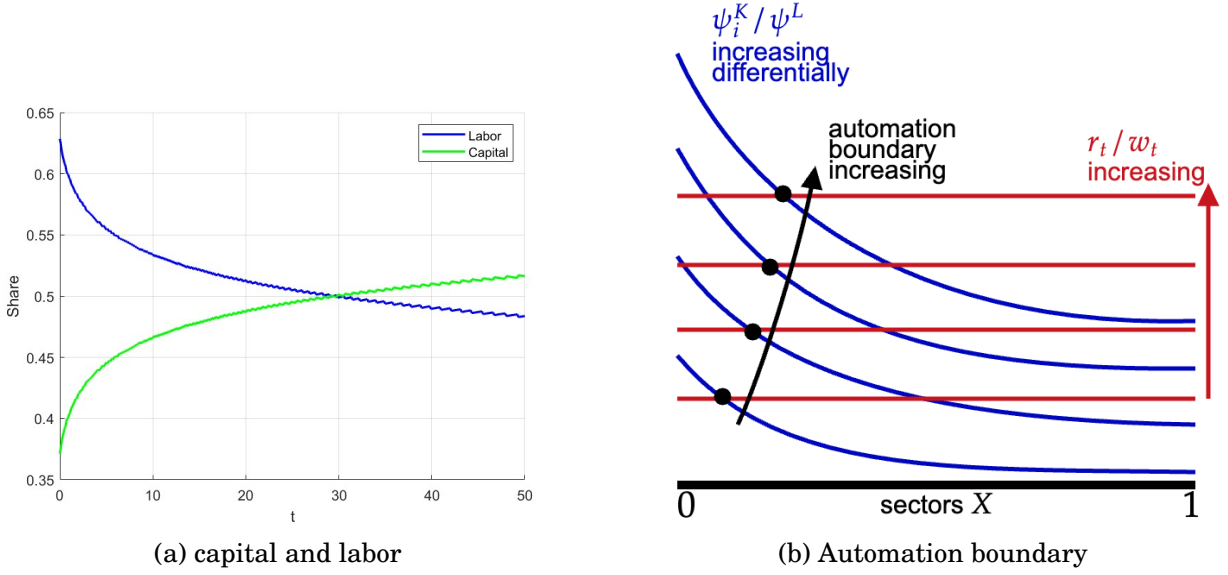
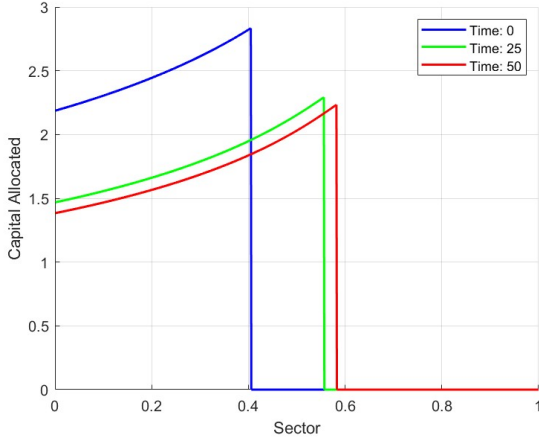


Figure 3: Dynamics of equilibrium outcomes when the degree of substitutability across tasks is low ($\sigma < 1/\eta$). In this case, capital allocation across tasks is biased toward data-poor tasks.

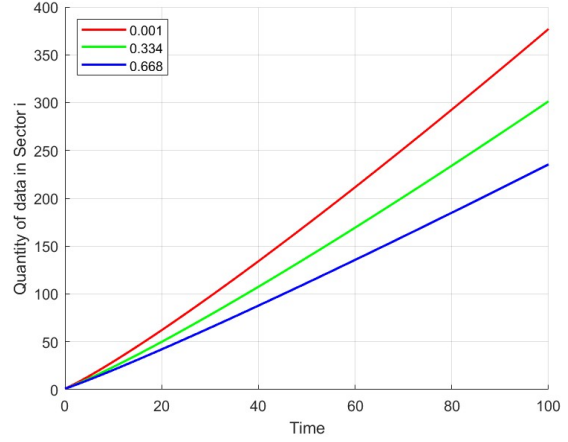
Next, consider the case in which tasks are gross substitutes, but smaller than the reciprocal of the degree of diminishing returns of data in capital productivity ($\sigma \in [1, 1/\eta)$). This case is more nuanced since the direction of capital allocation flips towards more productive tasks, so data growth might no longer be balanced—in principle, data-rich tasks can attract more capital, accumulate more data, and thereby attracting even more capital such that automation is plausibly constrained to a small set of highly capital productive tasks. We show that this cannot happen as the data feedback loop is dominated by diminishing returns from data: suppose again that the automation boundary is interior in the limit. Then, the ratio of productivity between the most and least productive task is *bounded* due to diminishing returns from data accumulation since $\eta < 1$, which immediately implies that the relative price of capital will remain finite as well. A similar argument therefore yields

$$\underbrace{\psi_{\gamma_t}^K}_{\rightarrow +\infty} = \underbrace{(r_t/w_t)}_{\rightarrow C < +\infty} \cdot \psi^L$$

yielding the same contradiction as before. The proof in [Appendix A](#) makes this heuristic argument formal. Moreover, like the gross complements case, full limit automation implies production will become entirely capital driven. In this case, however, extensive automation and intensive capital improvements both push toward a larger capital share given that tasks are gross substitutes.



(a) Capital biased towards data-poor tasks



(b) Data stock for different tasks

Figure 4: Balanced data growth

Note: We set $\psi^L = L = K = 1$, $D_0 := \mathbf{1}$, $f(i) = 1 - i$, $\eta = 0.2$, $g = \delta_{\{0\}}$, $\sigma = 0.5$.
Details of procedure in [Appendix C](#).

Figure 3b illustrates the evolution of automation boundary in this case. Red lines represent the relative factor price of capital and blue curves reflect the cross-sectional distribution of task productivities. The important observation is that the automation boundary γ_t monotonically increases, until the economy is fully automated in the limit. In this limit, capital share converges to unity. This implies that when tasks are gross complements or weak substitutes, the expansion of the automation boundary dominates the Baumol effect that stems from capital augmenting productivity and drives the growth in capital share.

3.2 High Substitutability Across Tasks

We now show that the automation boundary is tight, i.e., when $\sigma > 1/\eta$, automation can fizzle out. In this case, the economy is not fully automated in the limit and the growth path of data remains imbalanced. Of course, this can only happen if tasks and their initial data allocation are not ex-ante symmetric. Otherwise, since the CES aggregator is symmetric, if both initial data D_{i0} and task-specific data productivities f_i are constant across tasks, all tasks produce the same quantity along this ‘knife edge’ equilibrium path.²⁵

Definition 3 provides a sufficient condition for data growth to be imbalanced when the tasks are not ex-ante perfectly symmetric. This sufficient condition qualifies the

²⁵No matter how we break ties in order to choose which tasks to automate and which tasks to produce with labor.

degree of heterogeneity in task-specific data productivity, f_i .

Definition 3 (σ -regular). $\mathbf{f} : X \rightarrow \mathbb{R}_{>0}$ is σ -regular if there exists $L_{\sigma, \mathbf{f}} \in (0, 1)$ s.t.

$$\frac{\sigma}{\sigma-1} f(i) < \frac{1}{i} \int_0^i f(j) dj \quad \text{for all } i \geq L_{\sigma, \mathbf{f}}.$$

Let $\mathcal{F}(\sigma)$ be the set of $\mathbf{f}$ that are σ -regular.

σ -regularity states that there must be sufficient heterogeneity in how efficiently different tasks use data. It essentially requires the reverse hazard rate of f_i to fall faster than the reciprocal of the task index i^{-1} :

$$\underbrace{\frac{f(i)}{\int_0^i f(j) dj}}_{\text{'reverse hazard rate'}} < \frac{1}{i} \cdot \frac{\sigma-1}{\sigma}$$

This condition guarantees enough heterogeneity in how tasks use data so that the positive data feedback loop mechanism can take hold. Note that for a fixed vector $\mathbf{f}$, increasing task substitutability makes the σ -regularity condition more likely to be satisfied. Indeed, in the limit where $\sigma \rightarrow \infty$, even a near-constant $\mathbf{f}$ will result in interior automation. **Example 1** illustrates an instance of a σ -regular vector $\mathbf{f}$.

Example 1 (Example of σ -regularity). An example parameterization of $\mathbf{f}$ is $f(i) = 1 + \epsilon - i$ for any $\epsilon < (\sigma - 1)/2$. In this case, the condition for σ -regularity becomes

$$\frac{\sigma}{\sigma-1} (1 + \epsilon - j) < 1 + \epsilon - \frac{1}{2} j$$

so any $L_{\sigma, \mathbf{f}} \in (\frac{2+2\epsilon}{1+\sigma}, 1)$ suffices.

With the definition of σ -regularity at hand, we are ready to state our second main result, pertaining to bounded automation.

Proposition 2 (Bounded automation and imbalanced data). If $\sigma > 1/\eta$ and $\mathbf{f}$ is σ -regular, then:

- (i) **Partial automation:** The economy is never fully automated: $\lim_{t \rightarrow \infty} \gamma_t < 1$;
- (ii) **Imbalanced data:** For any task $j \in \text{Int}(X)$ that is either always produced using capital or always produced using labor after some finite time:

$$\lim_{t \rightarrow \infty} \frac{D_{0t}}{D_{jt}} = \infty.$$

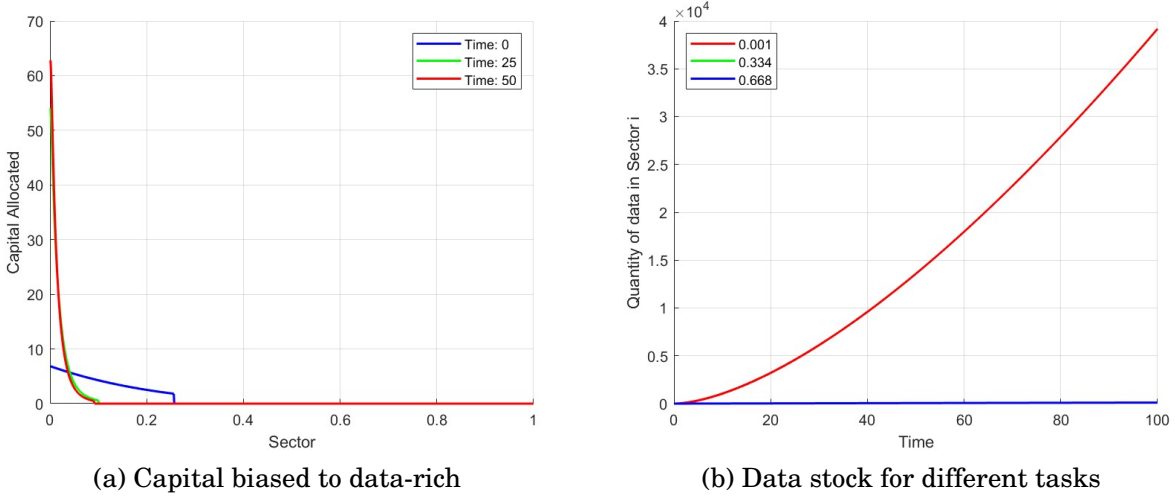


Figure 5: Imbalanced data growth
Note: parameters same as Figure 4 but with $\sigma = 5.5$.

Figure 5 illustrates Proposition 2. Panel 5a shows the evolution of capital allocation over time. There is a bias in capital allocation towards data-rich tasks, which in turn leads to a contraction of the automation frontier. Panel 5b plots the evolution of effective data stock of various tasks and illustrates that data becomes highly concentrated in a few capital intensive tasks.

The key intuition behind this result is that sufficient substitutability distorts the direction of capital allocation toward tasks that have high capital productivity. This, in turn, induces greater production of those tasks in general equilibrium, generating more task-specific data and amplifying the capital productivities of those tasks even further. This force pushes our economy toward *imbalanced data accumulation* and *partial automation*.

To see this, consider a similar thought experiment as in the previous section, where we run the economy forward in time, holding the automation boundary fixed. The relative rental rate can be written as:

$$\frac{r_t}{w_t} = \frac{L}{K} \cdot \frac{\psi^L}{1 - \gamma_t} \cdot \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\overbrace{\sigma - 1}^{>1}} dj = \frac{\psi_{\gamma_t}^K}{\psi^L}. \quad (11)$$

That is, the rental rate of capital relative to labor is a ‘weighted average’ of capital productivities among the inframarginally automated tasks. Holding the level of automation fixed, in this case we show the relative productivity growth of inframarginal tasks outpace that of the marginal task, a direct consequence of capital allocation

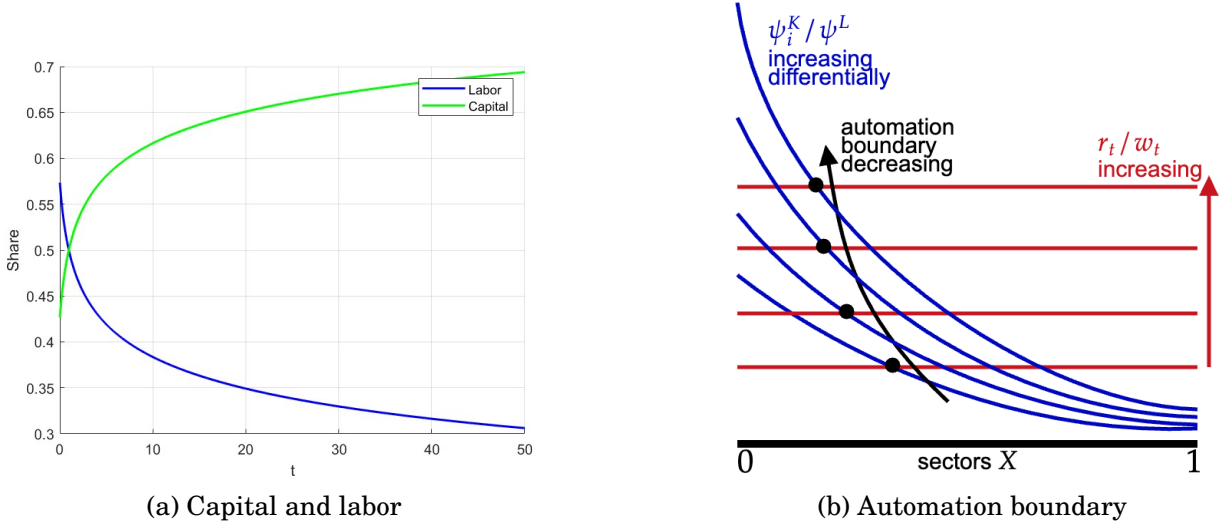


Figure 6: Dynamics of equilibrium outcomes when tasks are highly substitutable ($\sigma > \frac{1}{\eta}$). In this case, capital allocation across tasks is biased toward data-rich tasks.

being sufficiently biased towards data-rich tasks, so much so that it overcomes diminishing returns from data accumulation.²⁶ As such, γ_t would have to decrease to match the increased productivity of inframarginal tasks.

Figure 6b depicts the evolution of automation boundary when the degree of substitutability among tasks is sufficiently high, $\sigma > \frac{1}{\eta}$. Similar to the case of $\sigma < \frac{1}{\eta}$, red lines (blue curves) represent the relative factor price of capital (cross-sectional distribution of task productivities). In this case, the automation boundary γ_t must decline, for factor markets to clear, and the upper bound of automation boundary represent the automation level at which this occurs. As in the other case (gross complements and weak gross substitutes), capital share converges to one in the limit since $\lim_{t \rightarrow \infty} \inf_{\gamma \in [0,1]} \psi_{\gamma,t}^K = \infty$. We further verify numerically that the capital share increases along the transition path in Figure 6a. The driving force of the growth of capital share, however, is distinct from the low substitutability case. Since tasks are now gross substitutes, an increase in the capital index increases the capital share, whereas any resulting contraction in the automation boundary γ_t tames the initial increase. While in the gross complements case, the expansion of the automation boundary dominates the Baumol effect, here the increase in capital share suggests the opposite must be true.

Figure 7 summarizes Propositions 1 and 2. Panel 7a plots the evolution of the automation boundary, while Panel 7b plots the evolution of data used in task 0.5 (the task with median productivity) relative to task 0 (most productive task) over time,

²⁶Technically this only occurs in the regime where $\gamma_t \geq L_{\sigma,f}$. See Appendix A for additional details.

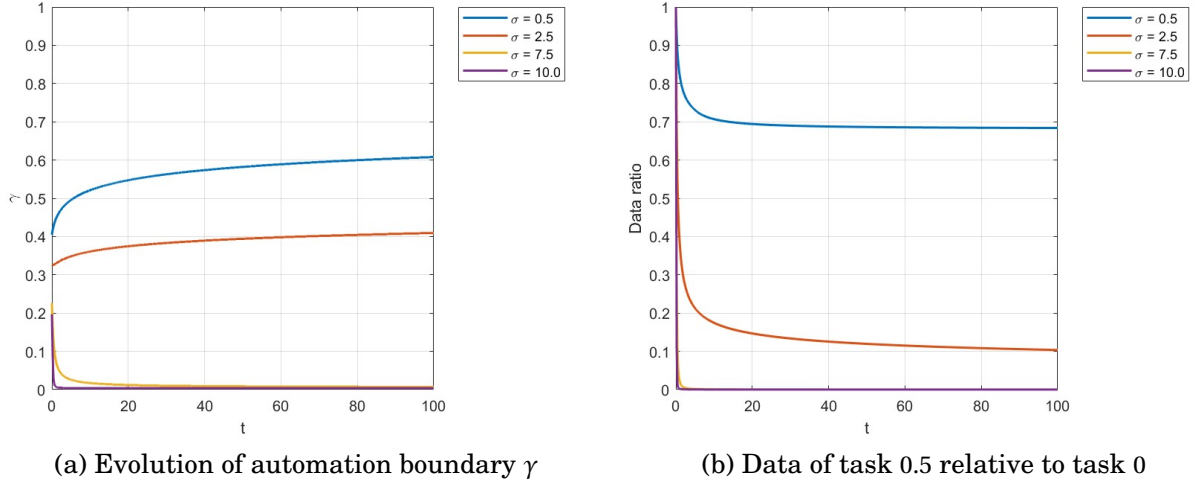


Figure 7: Comparative dynamics over time for different values of σ , when $\eta = 0.2$.

each for four different degrees of cross-task substitutability σ . We assume $\eta = 0.2$ such that $\sigma = 1/\eta = 5$ is the threshold that determines whether the economy is fully automated or not.

3.3 Wages

Taken together, our results on how task substitutability shapes the direction and limit of equilibrium automation has sharp implications for wages. When tasks are not strong enough substitutes (equivalently, sufficiently complementary) the expression of equilibrium wage in [Equation \(9\)](#) illustrates that there are two distinct effects of data-driven automation on the level of wages as in [Acemoglu and Restrepo \(2018\)](#). On the one hand, data-driven increases in capital productivity drives up the productivity of labor, thereby increasing wages. On the other hand, the contraction of labor onto a smaller set of tasks increases the effective labor supply, thereby reducing wages. The latter displacement effect is identical to [Acemoglu and Restrepo \(2018\)](#) but the former productivity effect stems from the role of data in augmenting capital productivity on the intensive margin rather than via the extensive margin by replacing inefficient labor. In the long-run they cancel out *exactly* such that wages eventually stagnate.

Alternatively, when tasks are sufficiently substitutable, [Proposition 2](#) implies that the automation boundary can be bounded above, thereby neutering the displacement effect that had offset wage growth in the complements case. As such, long-run wage growth will now be dominated by the productivity effect from data-driven automation. The following proposition formalizes this discussion.

Proposition 3 (Long-run wages). For any initial stock of data D_0 consistent with the

hypotheses of [Lemma 1](#):

(i) If $\sigma \leq 1/\eta$, then wages stagnate:

$$\lim_{t \rightarrow \infty} w_t = C < +\infty.$$

(ii) If $\sigma > 1/\eta$ and f is σ -regular, then wages grow unboundedly:

$$\lim_{t \rightarrow \infty} w_t = +\infty.$$

[Proposition 3](#) is tightly connected to the literature on how automation technologies impact wages in opposing direction. On the one hand, they can increase wages—by increasing the marginal productivity of labor. On the other hand, they can decrease wages by displacing the equilibrium set of tasks performed by labor (see, e.g., [Propositions 1 and 6 of Acemoglu \(2024\)](#)). [Proposition 3](#) makes precise how the endogenous composition of data accumulation and the ensuing direction of data-driven automation determine which of these forces dominate and when they offset each other exactly.²⁷ Indeed, we emphasize that [Proposition 4](#) is driven distinctly by endogenous data accumulation—and hence capital productivity: it is precisely when tasks are *complementary* that long-run wages stagnate because it drives the economy to produce more ‘data-poor’ tasks (e.g., that produced by labor). While this might aid wages in the short-run, this eventually augments the productivity of AI to perform such tasks. It is this sense that labor is *training their own replacement*.

Finally, it is worth noting that the wage stagnation result will hold when there is capital accumulation as well, as highlighted in the proof of [Proposition 3](#). Contemporaneous work by [Restrepo \(2025\)](#) similarly shows that wage will stagnate in an economy featuring capital accumulation but with fixed capital productivities across tasks. Allowing capital productivities to expand across tasks should increase the relative price of labor under gross complementarity. However, as argued earlier, the resulting displacement is sufficiently strong in our setting so as to keep wages constant. In this sense our result can be viewed as a strengthening of the wage stagnation proposition in [Restrepo \(2025\)](#).

²⁷This is another way in which data accumulation is distinct from capital accumulation: supposing our economy is fully automated in the limit ($\lim_t \gamma_t = 1$), long-run wages are, perhaps unsurprisingly, proportional to $K^{\frac{\sigma-1}{\sigma}}$ which is strictly decreasing in K whenever tasks are strictly gross substitutes ($\sigma < 1$), and strictly increasing in K .

3.4 Speed of Automation

We have thus far analyzed the limiting behavior of data-automation feedback loops. This raises a natural question of how *quickly* data-driven automation might take place. We offer a tight characterization:

Proposition 4 (Speed of data-driven automation). Suppose that $\sigma < 1/\eta$ and normalize $\psi^L = 1$. Then there exists some finite time $T < +\infty$ such that for all times $t \geq T$:

$$\frac{LM}{Kf} \left(\underline{B}^{1-\eta} + (1-\eta) \frac{K\bar{f}}{\underline{M}} (t-T) \right)^{-\frac{\eta}{1-\eta}} \leq 1 - \gamma_t \leq \frac{L\bar{M}}{K\underline{f}} \left(\underline{B}^{1-\eta} + (1-\eta) \frac{Kf}{\bar{M}} (t-T) \right)^{-\frac{\eta}{1-\eta}} \quad (\text{S})$$

where $\underline{M}, \bar{M}, \underline{B}, \bar{B} > 0$ are constants, $\underline{f} := \min f$, and $\bar{f} := \max f$. Importantly, the fraction of tasks produced by labor decays as a power law, $1 - \gamma_t = \Theta\left(t^{-\frac{\eta}{1-\eta}}\right)$. Moreover, if $\mathbf{D}_0$ is less imbalanced than the the balanced-data limit, i.e.,

$$\frac{D_{i0}}{D_{j0}} \leq \left(\frac{f_i}{f_j} \right)^{\frac{\sigma}{1-\sigma\eta}} \quad \forall i < j,$$

then the above bounds on speed (S) holds for all times $t \geq 0$ (i.e., setting $T = 0$).

Proposition 4 states that even when data-driven automation leads to full limit automation, equilibrium automation can be surprisingly slow—the share of tasks produced by labor decays as a power law in time.

Note that the first part of **Proposition 4** is stated as an asymptotic result, so it holds for any initial data stock. Alternatively, the second part imposes an additional condition on the initial data stock, that it is not more imbalanced than the long-run composition (as given by **Proposition 1**), which then delivers the same bounds on equilibrium non-asymptotically i.e., for all times $t \geq 0$.

This allows us to sharpen our analysis of the speed of equilibrium automation along the transition path. The bound on the process in (S) has the generic form

$$1 - \gamma_t \simeq A \left(B^{1-\eta} + Ct \right)^{-\frac{\eta}{1-\eta}} = AB^{-\eta} \left(1 + \frac{t}{\tau} \right)^{-\frac{\eta}{1-\eta}} \quad \text{where } \tau := \frac{B^{1-\eta}}{C}$$

and $A, B, C > 0$ are constants independent of time. Here B can be interpreted as the relevant local data stock at date 0, while C summarizes the effective rate at which accumulated production expands the data stock near the automation boundary.²⁸ Dif-

²⁸For the lower bound on $1 - \gamma_t$, $B = \bar{B}$ and $C = (1-\eta)K\bar{f}/\underline{M}$. For the upper bound on $1 - \gamma_t$, $B = \underline{B}$ and $C = (1-\eta)Kf/\bar{M}$.

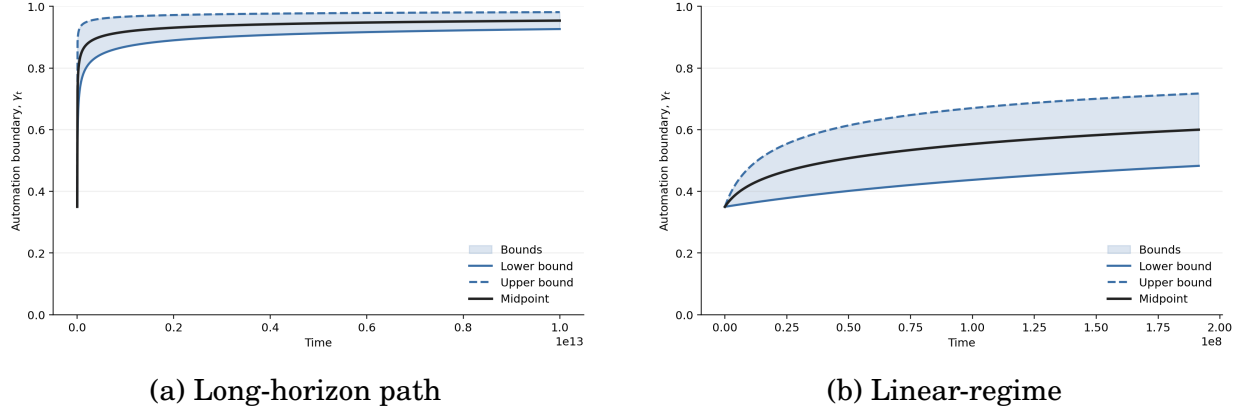


Figure 8: Bounds on the speed of data-driven automation

ferentiating the generic process gives

$$\dot{\gamma}_t = \frac{\eta}{1-\eta} \frac{AC}{B} \left(1 + \frac{t}{\tau}\right)^{-\frac{1}{1-\eta}}.$$

This expression showcases two speed regimes: When $t \ll \tau$, the inherited data term $B^{1-\eta}$ dominates the accumulated-flow term $C \times t$:

$$\dot{\gamma}_t \simeq \frac{\eta}{1-\eta} \frac{AC}{B}.$$

Hence the automation boundary grows linearly in time early in the process since initial phase is governed by the inherited data stock near the automation boundary. By contrast, when $t \gg \tau$, the accumulated-flow term $C \times t$ dominates and

$$\dot{\gamma}_t \simeq \frac{\eta}{1-\eta} AC^{-\frac{\eta}{1-\eta}} t^{-\frac{1}{1-\eta}}.$$

which slows down over time because of diminishing returns to data accumulation.

Section 3.4 illustrates this by plotting bounds for the automation boundary γ_t .²⁹ Panel (a) plots the path of automation over a long horizon. Note here that we see a ‘sudden’ increase in automation (as measured on that time scale) followed by a tapering off: when $\approx 95\%$ of tasks are already automated, the speed of automation progressively slows. Panel (b) zooms in on the early transition by a factor of 10^5 : here the path is approximately linear in calendar time.

²⁹We set $\eta = 0.2$, $\sigma = 0.5$, $K = 1$, $L = 120$, $\underline{f} = 0.95$, $\bar{f} = 1$, and $D_{i0} = 10^{10}$ for all tasks.

4 Contagious Automation

In this section, we introduce cross-task spillovers into our economy as captured by the graphon $W : [0, 1]^2 \rightarrow [0, 1]$ and analyze how it shapes (i) whether the economy is fully automated in the limit; (ii) the long-run composition of data and production; and (iii) the speed of automation.

4.1 Limit Automation with Spillovers

Connectedness implies full automation. With cross-task spillovers, automation will not, in general, take a threshold form since a high index task might, through spillovers, accumulate a higher effective data stock than a lower index task. To this end, we will define:

$$\mathcal{K}_t := \left\{ i \in X : \frac{r_t}{\psi_{i,t}^K} \leq \frac{w_t}{\psi^L} \right\} \quad \text{and} \quad \mathcal{L}_t := X \setminus \mathcal{K}_t.$$

as the set automated tasks and tasks produced by labor at time t . We let μ denote the Lebesgue measure so full automation means $\mu(\mathcal{L}_t) \rightarrow 0$.

Definition 4 (Uniform strong connectedness). A directed graphon W is uniformly strongly connected if there exist $n \in \mathbb{N}$ and $\varepsilon > 0$ such that

$$W^{(n)}(i, j) \geq \varepsilon \quad \text{for a.e. } (i, j) \in X^2.$$

where $W^{(n)}$ is defined as follows: $W^{(1)} = W$ and

$$W^{(m+1)}(i, j) := \int_X W(i, k) W^{(m)}(k, j) dk \quad \text{for } m \geq 1.$$

Thus $W^{(m)}(i, j)$ is the total intensity of length- m directed spillover paths from source j to beneficiary i .

We call this uniform strong connectedness because every task can receive data, directly or indirectly, from every other task within a uniformly bounded number of spillover steps and with uniformly positive intensity.

Proposition 5 (Contagious automation). Suppose W satisfies uniform strong connectedness. For any $\sigma > 0$, any $\eta \in (0, 1)$, and any bounded initial data stock $\mathbf{D}_0 \in L_+^1(X)$, every equilibrium path satisfies full automation:

$$\mu(\mathcal{L}_t) \rightarrow 0.$$

Proposition 5 states that connectedness *per se* is sufficient for full limit automation, even if cross-task production is very substitutable ($\sigma \gg 1$) and there are enormous differences in returns to data across tasks (f_i is steeply decreasing in i). The basic intuition is that when tasks are highly substitutable ($\sigma > 1/\eta$) *local* spillovers that generate a connected graphon can stand-in for *global equilibrium* forces from the gross-complements case ($\sigma \leq 1$) that guides capital allocation and thereby data accumulation toward data-poor tasks.³⁰

Put differently, the increasing return to scale that follows any degree of data spillover across tasks is so strong that it overcomes any force towards capital concentration, caused by high degree of task substitutability. As such, the degree of task complementarity that is required to distribute capital across tasks when the economy is in data autarky is not needed when there is data spillover across tasks.

That local spillovers *per se* are sufficient is perhaps surprising, and goes against the broad intuition that connectedness is often insufficient for well-behaved limit behavior in networks with nonlinear dynamics. For example, in models of opinion dynamics, it is well-known that connectedness *per se* is not enough for convergence under non-linear aggregation of opinions (Cerreia-Vioglio, Corrao, and Lanzani, 2024) or in binary-action coordination games (Morris, 2000). Our setting is distinct from these models because (i) equilibrium factor prices (a ‘global’ variable) shape the path of data accumulation $(\mathbf{D}_t)_t$; (ii) our state space is unbounded; and (iii) in the long-run the whole vector of data and productivities diverge to infinity ($\mathbf{D}_t \rightarrow +\infty$ and $\boldsymbol{\psi}_t^K \rightarrow +\infty$) whether or not the graphon is connected.

A key feature governing whether automation is contagious or not is the *limit ratio of effective task-specific data*, which—for any two tasks $i, j \in \text{Int}(X)$ —is given by

$$\lim_{t \rightarrow \infty} \frac{\mathcal{A}_i(\mathbf{D}_t)}{\mathcal{A}_j(\mathbf{D}_t)} =: \mathcal{R}_{ij}.$$

A sufficient condition for contagious automation is that this ratio does not go to either 0 or infinity i.e., $0 < \mathcal{R}_{ij} < +\infty$. If it did, then in the case where tasks are arbitrarily substitutable ($\sigma \gg 1$), one of those tasks need not be produced by capital in equilibrium.

What governs the ratio of effective task-specific data $\mathcal{R}_{ij}$? A loose intuition is that connectedness ensures for some task $k \in X$ close to i , their neighbors overlap non-trivially, as illustrated by **Figure 9** that illustrates a graphon W in which tasks that are closer to each other (in index) have stronger spillovers. But this, *per se*, does not imply that the limit ratio of effective data $\mathcal{R}_{ik}$ is bounded away from 0 because their

³⁰This overturns **Proposition 2**, which establishes partial automation in a data autarky when tasks are sufficiently substitutable.

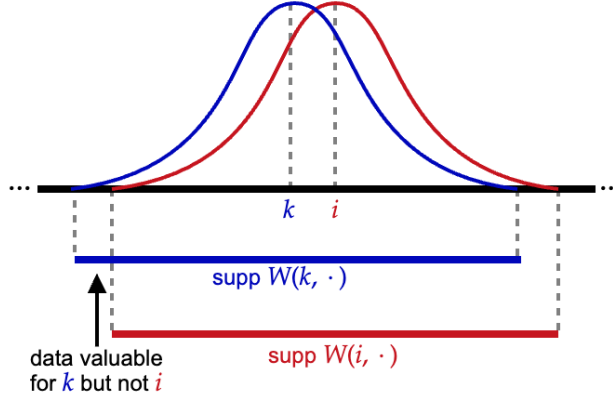


Figure 9: Overlapping neighbors

supports are non-overlapping—task k but not i could value the data of some other task $\ell \in X$. Then, if $D_{\ell t}$ grows quickly, the limit ratio of effective data $\mathcal{R}_{ik}$ could, in principle, be 0. Nonetheless in [Appendix A.2](#) we show via a chaining argument that, because the space of tasks X is bounded, this cannot happen *in equilibrium*—connectedness implies that along the equilibrium path full automation is inevitable.³¹

4.2 How Do Spillovers Shape Speed?

We next numerically investigate the speed of automation in the presence of spillovers. [Figure 10](#) illustrates the impact of varying the breadth and depth of cross-task spillovers on the evolution of the automation boundary, separately for the gross complements case and the gross substitutes case.

In the case of complements, automation increases roughly linearly in log time with or without spillovers. Intensifying the *breadth* of spillovers while keeping the intensity of spillovers constant has little impact on the pace of automation. This is because data across tasks are, in equilibrium, already balanced absent spillovers per [Proposition 1](#), so the benefit from being able to draw on other tasks’ data is muted. If, however, we intensify the *depth* of spillovers i.e., increasing the intensity of spillovers, this generates a ‘one-time’ boost to automation since doing so is equivalent to increasing the degree to which each task values data. This is illustrated in Panel [10a](#).

In the case of substitutes, whereas automation declines linearly in log time absent spillovers, the trajectory flips when spillovers are introduced, consistent with [Proposition 5](#). Indeed, we know data will be highly skewed towards data-rich tasks in the case where $\sigma > 1/\eta$, so the ability to draw on their data has a dramatic impact on

³¹In [Appendix A.2.1](#) we also characterize the long-run composition of task production and data as a function of the leading eigenfunction of the graphon.

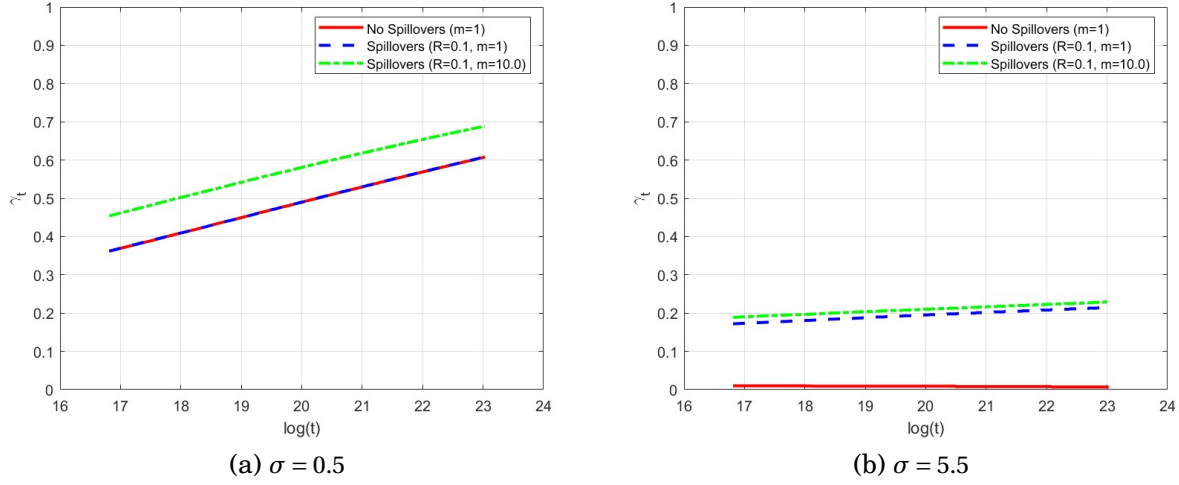


Figure 10: Evolution of γ under alternative spillover geometries

Note: m denotes the measure of the $g(\cdot, \cdot)$ kernel that induces $W(\cdot, \cdot)$, which captures the depth of spillovers. The remaining parameters are as in Figure 4 except with $L = 100$.

furthering automation. Scaling the depth of spillovers similarly increases automation, but the impact is less pronounced compared to scaling breadth in the case of substitutes in the presence of imbalanced data. This is illustrated in Panel 10b.

The broad economic takeaway from this exercise is that equilibrium automation with and without spillovers is surprisingly slow—approximately linear growth in *log-time*. This is consistent with Proposition 4 that shows that in the absence of spillovers, the automation boundary exhibits ‘power law’ behavior in time with a ‘fat tail’ of tasks performed by labor. Here, data spillovers lead to a *level change* in the automation boundary, but does not change its long run speed. Of course, this analysis

Throughout the analysis above, we have kept the capital stock fixed. This might approximate the shorter run in which supply chain constraints prevent rapid scaling up of computational resources. The takeaway is that while data-driven automation is powerful on its own, it is unlikely to speed up automation substantially in the long run. But in the short-run, we will see that the network structure plays an important role in shaping automation dynamics.

4.3 Core-Periphery Automation

We now illustrate the role of network structure in short run dynamics via the following stylized example. Partition tasks $X = [0, 1]$ into a core C of measure α and a periphery P of measure $1 - \alpha$. Tasks are homogeneous within each block: $f_C > f_P$ and

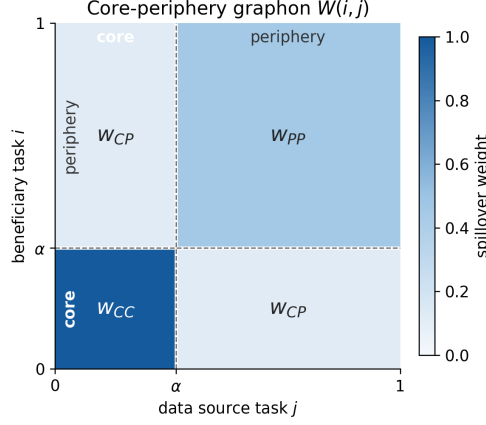


Figure 11: A two-block core-periphery graphon. Rows index the beneficiary task i , columns index the data source task j , and darker shading indicates stronger spillovers.

$D_{C0} > D_{P0}$. Effective data at the block level is

$$A_C = \underbrace{w_{CC}D_C}_{\text{core} \leftrightarrow \text{core}} + \underbrace{w_{PC}D_P}_{\text{periphery} \rightarrow \text{core}} \quad \text{and} \quad A_P = \underbrace{w_{CP}D_C}_{\text{core} \rightarrow \text{periphery}} + \underbrace{w_{PP}D_P}_{\text{periphery} \leftrightarrow \text{periphery}}.$$

Capital productivities are $\psi_C = f_C A_C^\eta$ for core tasks and $\psi_P = f_P A_P^\eta$ for periphery tasks. Our economy evolves across two phases. In the first phase, only core sectors are automated. In the second phase, automation begins to take hold in the wider economy. There are two crucial forces at play:

- *Core data-feedback loops*: the overproduction of core sectors (relative to peripheral sectors) generate data that exhibit strong within-block spillovers (high w_{CC}). This drives up the requisite productivity required for non-peripheral sectors to be automated: the opportunity cost of deploying capital in peripheral sectors goes up when core sectors are productive.
- *Bottlenecked spillovers*: If there are only weak spillovers from core sectors e.g., software and coding related tasks to peripheral sectors (small w_{CP}). This means that the capital productivity of peripheral sectors lag behind, and it takes longer for automation to reach the wider economy.

Both of these forces can be gleaned from the following equation that pins down the requisite degree of capital productivity in peripheral sectors for automation to begin in the wider economy:

$$\left(\frac{L\alpha}{K(1-\alpha)} \right)^{1/\sigma} \psi_C^{\frac{\sigma-1}{\sigma}}.$$

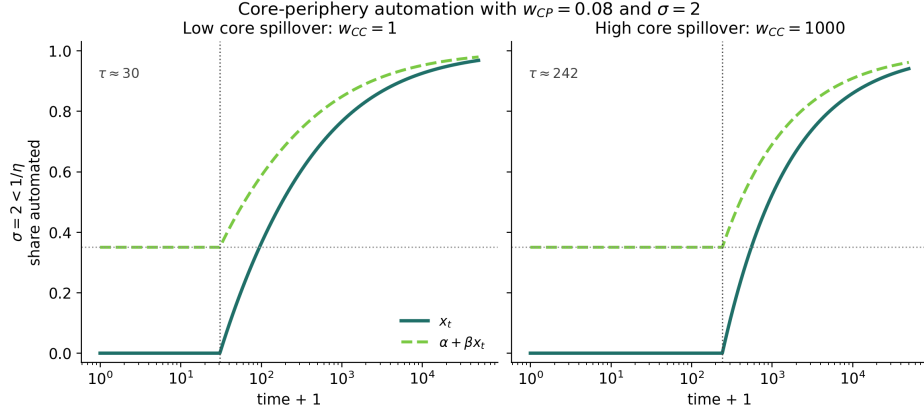


Figure 12: Automation paths

and the time τ at which the wider economy begins to be automated can be analytically characterized via a pair of coupled ODEs. We emphasize here that the automation can have a non-monotonic relationship with link strength: by intensifying the degree of within-core spillovers w_{CC} , this can delay the time at which peripheral tasks begin to be automated. This is because ψ_C now grows more quickly which, in turn, sucks up capital. The corresponding automation paths are illustrated in [Figure 12](#) where the dotted green lines show the total share of sectors automated in the economy, while the blue line illustrates the total share of the periphery automated. The left panel shows the case with low intensity spillovers within the core, and the right panel shows the case with high intensity spillovers.³²

Taken together, our results suggest that although the long-run speed of the economy exhibits power law behavior, the short-run dynamics can depend richly on the network structure of data spillovers.

5 Is Data-Driven Automation Efficient?

We have thus far offered an analysis of when and why our economy exhibits full limit automation, with and without data spillovers. A natural question is whether this degree of automation is efficient.

A basic observation is that because all producers in our economy are atomless and data generated from economic activity of each producer is shared for free among all intermediate good producers, they do not internalize the value of producing more data today on future production—both for their own task or for other tasks that might benefit from their data. Indeed, modern machine learning systems rely on vast amounts

³²A fuller set of simulations for different parameter values is in [Appendix D.2](#).

of data that no single firm produces on their own so they might not internalize this externality in equilibrium.

In this section, we again restrict attention to the economy in data autarky and compare the equilibrium path of data accumulation with the first-best benchmark. To preview our results—under two simplifications to be discussed—we show that the myopic equilibrium path is generically suboptimal. Specifically, the planner would amplify the market’s bias: in the gross complements and weak substitutes case ($\sigma \leq 1/\eta$), the planner would allocate even more capital to tasks with less data to speed up the data convergence exhibited in the decentralized equilibrium; in the strong substitutes case ($\sigma > 1/\eta$), the planner would allocate even more capital to tasks with more data, thereby accelerating the data divergence observed in the decentralized equilibrium.

Simplifying assumptions: The coarsened economy. To focus on planner’s incentives to ‘tilt’ the direction of capital allocation, we study a simpler, coarsened economy that is more amenable to analysis. We make two simplifications of our main model (in data autarky) to this end. First, we suppose production is entirely capital-driven i.e., $L = 0$ and we are in an ‘AK-style’ economy, albeit with no capital accumulation. This simplification allows us to focus on the direction of capital tilting; we expect a similar mechanism would be at play for the allocation of labor. Second, we collapse our tasks into two blocks for tractability. We call the task-block $[0, 1/2]$ ‘ h ’ (high data productivity), and the task-block $(1/2, 1]$ ‘ ℓ ’ (low). Qualitative insights from this simple two task-block case should generalize to our baseline economy featuring a continuum of tasks.

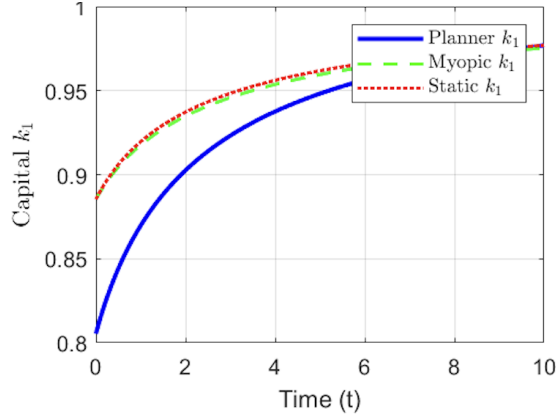
Planner’s problem. Consider a planner who chooses the allocation of capital to maximize output dynamically for some discount rate $\rho > 0$. The planner solves:

$$\begin{aligned} \max_{k_{1t}, k_{2t}} \int_0^T e^{-\rho t} \underbrace{\left(\sum_{i \in \{h, \ell\}} \frac{1}{2} (\psi^K(D_{it}) k_{it})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}}_{\equiv Y_t} dt \quad (12) \\ \text{s.t.} \quad \sum_{i \in \{h, \ell\}} \frac{1}{2} \cdot k_{it} = K \quad \text{and} \quad D_t = D_0 + \int_0^t y_s ds \end{aligned}$$

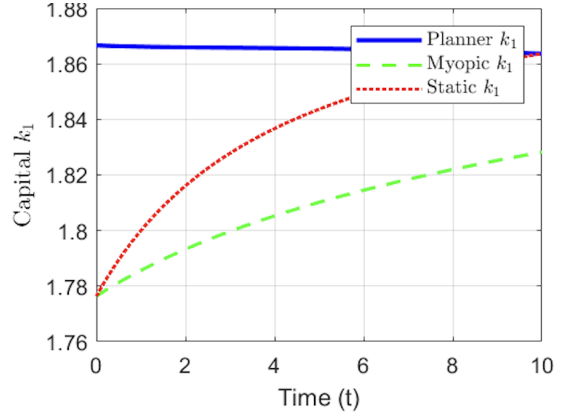
where the constraints are the capital market clearing condition and the usual data stock law of motion.³³ We say that the equilibrium is efficient if it solves the planner’s problem (12). **Proposition 6** summarizes our main efficiency result.

Proposition 6 (Inefficiency and planner’s problem). If $f_h = f_\ell$ and $D_{h0} > D_{\ell 0}$, then the equilibrium path is efficient iff $\sigma = 1/\eta$. Otherwise:

³³That is, $D_{it} = D_{i0} + \int_0^t y_{is} ds$ for each $i \in X$.



(a) Capital allocated to task-block 1
when $\sigma = 0.5$



(b) Capital allocated to task-block 1
when $\sigma = 5.5$

Figure 13: Numerical illustration of the planner's solution

- (i) Compared to the myopic allocation given planner's current data stock, the planner allocates—for all $t \in [0, T]$ —weakly more capital to l if $\sigma < 1/\eta$ and weakly more capital to h otherwise.
- (ii) Compared to the *decentralized equilibrium*, the planner allocates
 1. weakly more capital to l initially, but may reverse if $\sigma < 1$,
 2. weakly more capital to l if $1 \leq \sigma < 1/\eta$, for all $t \in [0, T]$,
 3. weakly more capital to h if $\sigma > 1/\eta$, for all $t \in [0, T]$.

The difference between planner's decision how to allocate capital (unequally) compared to the market's decision is governed by the following inequality:

$$\frac{\lambda_{lt}}{\lambda_{ht}} \leq \frac{\psi_{ht}^K}{\psi_{lt}^K}, \quad (13)$$

where λ_{it} , $i \in \{l, h\}$ is the co-state variable capturing the value of block i 's data. Away from the knife edge case where $\sigma = 1/\eta$ where the above relation holds with equality at all times, the equilibrium path of capital allocation is generically inefficient.

The important observation is that the direction of inefficiency in equilibrium capital allocation depends on the degree of cross-task substitutability. If $\sigma < 1/\eta$, we show the left hand side of Equation (13) will be greater, or that the relative value of data for low data tasks is sufficiently high so as to justify distorting production presently. In this case, the market inefficiently allocates too much capital to tasks that are currently data-rich—even if the equilibrium already biases capital allocation toward the data-poor block ℓ , the planner prefers to tilt this allocation more aggressively. In so doing,

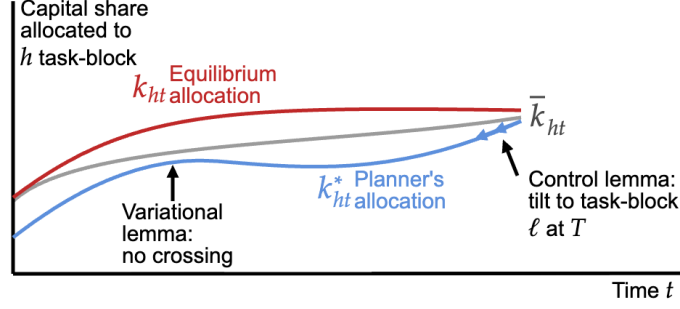


Figure 14: Illustration of proof ideas for $1 < \sigma < 1/\eta$

the planner speeds up the rate at which block ℓ 's data is accumulated—at the cost of static misallocation—in order to loosen the bottleneck more quickly.

Conversely, when tasks are highly substitutable $\sigma > 1/\eta$, the right hand side of [Equation \(13\)](#) will be larger, so the planner is incentivized to lean into tasks that are *already* data-rich—even though the equilibrium already biases capital allocation toward the data-rich block h , the planner prefers to tilt this allocation more aggressively. In doing so, the planner harnesses the data-feedback loop to her advantage. The proof of [Proposition 6](#) formalizes the above argument.

[Figure 14](#) illustrates the intuition for the case when $\sigma < 1/\eta$. Write $\mathbf{k}_h^*$ to denote the planner's optimal path of capital allocation to block h , and $\mathbf{k}_h$ to denote the equilibrium path. For each task-block h , $\mathbf{k}_h^*$ and $\mathbf{k}_h$ are two infinite-dimensional vectors, where the dimension is time. To show that $k_{ht}^* \leq k_{ht}$ for all t , as claimed by the statement of [Proposition 6](#), we combine a control argument to sign the direction of tilting in the long-run by running time backwards, and a variational argument to argue that the planner's optimal capital allocation cannot cross an auxiliary path of capital $\bar{\mathbf{k}}$ that maximizes static time- t production given the same data stock as the planner's.

Reintroducing labor to the coarsened economy. [Proposition 6](#) applies to an economy with no labor, which does not speak to the *speed* of automation under the planner's problem. In [Appendix D.4](#), we reintroduce labor into the economy when $\sigma \leq 1/\eta$ and show numerically that the planner, as before, continues to tilt capital towards the data-poor task. This, in turn, *accelerates* automation relative to the decentralized equilibrium. We expect the opposite to be true in the $\sigma > 1/\eta$ case, as the planner's amplified bias towards data-rich tasks concentrates the use of capital and thereby shrinks the frontier of automation.

6 Capital Accumulation

We have thus far analyzed an economy populated by hand-to-mouth households, without any capital accumulation. This was a deliberate modeling choice to focus on the role of heterogeneous data accumulation. Of course, capital accumulation—in the form of robotics, data centers, computational power, and so on—plays a central role in driving modern AI systems. We will accordingly augment our baseline model with capital accumulation but adopt the simplifying assumption that tasks are symmetric, under which the economy can be well approximated by an ‘AK’ representation, and rule out cross-task spillovers. This is, in some sense, the ‘most pessimistic’ case and will serve as a lowerbound on growth for an economy with higher substitutability and/or spillovers. We emphasize simplification is largely cosmetic and does not affect the core result that output and consumption become unbounded in finite time. That said, the symmetry assumption does have bite for the share of automated tasks; see the discussion of [Proposition 2](#) for details.

We now introduce the standard household’s consumption-savings problem. The representative household solves,

$$\max_{c_t, a_t} \int_0^\infty e^{-\rho t} \frac{c_t^{1-\theta}}{1-\theta} dt \quad s.t. \quad \dot{a}_t = w_t + r_t a_t - c_t$$

where as usual θ is the CRRA parameter, $\rho > 0$ is the discount rate, and a_t is the wealth at time- t which earns a risk-free rate of r_t . L is fixed and normalized to one. Our main result in this section is that data-driven automation combined with capital accumulation can deliver runaway growth (output becomes unbounded in finite time), which is reminiscent of the ‘supply-side singularity’ of [Nordhaus \(2021\)](#). We shall return to discuss two major departures at the end of the section.

In light of the possibility of runaway growth, we have to be more careful with what it means for households to be trading off consumptions and savings decisions optimally. We introduce the following definition:

Definition 5 (Feasibility and optimality of paths). A path $(c_t, a_t)_t$ is feasible if it satisfies the household budget constraint. A feasible path is optimal if either:

- (i) It delivers infinite present discounted utility.
- (ii) It satisfies both A and B:
 - A. The Euler condition: $\dot{c}_t/c_t = (r_t - \rho)/\theta \quad \forall t \geq 0$.
 - B. The transversality condition: $\lim_{t \rightarrow \infty} e^{-\rho T} a_t/c_t^\theta = 0$.

Absent capital accumulation, it is straightforward to verify that our economy features no growth in the limit akin to an ‘AK’ economy with a fixed level of capital produc-

tivity, with growth being positive but finite. By combining (i) data accumulation that runs into diminishing returns but does not depreciate and (ii) endogenous capital accumulation that does depreciate, we show our economy exhibits a singularity: output becomes unbounded in finite time.

Proposition 7 (The data-driven singularity). There exists a feasible and optimal path for the household’s problem such that by some finite time T , consumption and output are unbounded and the economy is fully automated. That is,

$$\lim_{t \rightarrow T} \gamma_t = 1, \quad \lim_{t \rightarrow T} \frac{d \ln Y_t}{dt} = +\infty, \quad \lim_{t \rightarrow T} \frac{d \ln c_t}{dt} = +\infty.$$

To develop intuition for [Proposition 7](#), consider an economy satisfying $Y_t = A_t K_t$ with an exogenous savings rate $s \in (0, 1)$ à la Solow. Then changes in output can be bounded below as follows:

$$\dot{Y}_t = \dot{A}_t K_t + A_t \dot{K}_t \geq A_t \dot{K}_t = (s - \delta / A_t) A_t Y_t$$

where the first inequality assumes $\dot{A}_t \geq 0$. If A_t is of the form Y_t^ω for some $\omega > 0$, then output growth is explosive. We establish our economy indeed features this constant elasticity relation between productivity and output in the proof.

[Proposition 7](#) echoes the ‘supply-side singularity’ of [Nordhaus \(2021\)](#) though there are two substantial differences. First, [Nordhaus \(2021\)](#) assumes the capital productivity term grows exogenously at some exponential rate; by contrast, data is (i) accumulated endogenously in our economy; and (ii) exhibits potentially steeply diminishing returns. The key is that data-accumulation and capital-accumulation are *strong complements*: either, on their own, can only generate finite growth, but they feed into each other to generate unbounded growth. Second, [Nordhaus \(2021\)](#) assumes that labor and capital are gross substitutes; [Trammell and Korinek \(2025\)](#) also note capital-driven growth typically requires gross substitutability across capital and labor. On the other hand, our result does not restrict σ to be greater than one. In our economy, for a *fixed* level of automation, we see from [\(4\)](#) that capital and labor may well be complementary, but as data and/or capital accumulates, the automation boundary endogenously expands so capital dominates production and (almost) obviates the role of labor. Put differently, whereas the complementarity between capital and labor at any moment in time is defined by task complementarity σ , data-driven productivity gains for capital will eventually push the task-level perfect substitutability to dominate.³⁴

³⁴Also related to this result is contemporaneous work by [Jones and Tonetti \(2025\)](#) who find a singularity through the lens of semi-endogenous growth model provided that capital benefits sufficiently from the idea stock.

7 Conclusion

In this paper, we develop a model of data-driven automation featuring (i) task-specific data creation; (ii) endogenous data accumulation; and (iii) cross-task data spillovers, in general equilibrium. These features generate rich equilibrium dynamics in which data plays a dual role in *augmenting* the productivity of already-automated tasks and *expanding* the automation frontier. We provided tight conditions on the degree of cross-task substitutability for the economy to be fully automated in the limit (**Propositions 1 and 2**), and established how cross-task complementarity can lead long-run wages to stagnate (**Proposition 3**). We further developed tight non-asymptotic bounds on the speed of equilibrium data-driven automation (**Proposition 4**), showing that it can be surprisingly slow with a ‘fat tail’ of tasks performed by humans even in the long run. We also proved that *any* degree of connectedness via cross-task data spillovers is sufficient to generate full limit automation (**Proposition 5**). Even so, equilibrium automation can be inefficiently slow with data externalities distorting the equilibrium composition of data accumulation (**Proposition 6**). Finally, we showed that the ability for data to simultaneously augment capital and expand the set of automated tasks can—even if capital runs into steeply diminishing returns from data—amplify capital accumulation to generate unbounded long-run growth (**Proposition 7**).

Our baseline model can be extended to include multiple sectors, feature non-separable data aggregators, and accommodate a time-varying set of tasks. These represent plausible avenues for future work. A more detailed discussion can be found in **Appendix B**.

References

- ACEMOGLU, D. (2002): “Directed technical change,” *The review of economic studies*, 69, 781–809.
- (2024): “Capital and wages,” Tech. rep., National Bureau of Economic Research.
- ACEMOGLU, D., F. KONG, AND P. RESTREPO (2025): “Tasks at work: comparative advantage, technology and labor demand,” *Handbook of Labor Economics*, 6, 1–114.
- ACEMOGLU, D. AND P. RESTREPO (2018): “The race between man and machine: Implications of technology for growth, factor shares, and employment,” *American economic review*, 108, 1488–1542.
- AGHION, P., B. F. JONES, AND C. I. JONES (2017): “Artificial intelligence and economic growth,” Tech. rep., National Bureau of Economic Research.

- ARMSTRONG, T. E. (1993/94): “Chebyshev Inequalities and Comonotonicity,” *Real Analysis Exchange*, 19, 266–268.
- ARROW, K. J. (1962): “The economic implications of learning by doing,” *The review of economic studies*, 29, 155–173.
- AUTOR, D. AND N. THOMPSON (2025): “Expertise,” *Journal of the European Economic Association*, jvaf023.
- AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): “The skill content of recent technological change: An empirical exploration,” *The Quarterly journal of economics*, 118, 1279–1333.
- AXIOS (2025): “AI jobs danger: Sleepwalking into a white-collar bloodbath,” *Axios*, interview with Dario Amodei, CEO of Anthropic, warning about AI-driven job losses and the risk of a “white-collar bloodbath”.
- BAJARI, P., V. CHERNOZHUKOV, A. HORTAÇSU, AND J. SUZUKI (2019): “The impact of big data on firm performance: An empirical investigation,” in *AEA papers and proceedings*, American Economic Association, vol. 109, 33–37.
- BAUMOL, W. J. (1967): “Macroeconomics of unbalanced growth: the anatomy of urban crisis,” *The American economic review*, 57, 415–426.
- BERAJA, M. AND N. ZORZI (2025): “Inefficient automation,” *Review of Economic Studies*, 92, 69–96.
- BESIROGLU, T., E. ERDIL, M. BARNETT, AND J. YOU (2024): “Chinchilla scaling: A replication attempt,” *arXiv preprint arXiv:2404.10102*.
- BORGS, C., J. T. CHAYES, L. LOVÁSZ, V. T. SÓS, AND K. VESZTERGOMBI (2008): “Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing,” *Advances in Mathematics*, 219, 1801–1851.
- CAPOOT, A. (2025): “Legal AI startup Harvey hits \$100 million in annual recurring revenue,” *CNBC*, accessed October 18, 2025.
- CERREIA-VIOGLIO, S., R. CORRAO, AND G. LANZANI (2024): “Dynamic opinion aggregation: long-run stability and disagreement,” *Review of Economic Studies*, 91, 1406–1447.
- CHATTERJI, A., D. ROCK, AND E. TALAMAS (2025): “Transformative AI and Firms,” Tech. rep., National Bureau of Economic Research, Inc.
- CHEN, J., M. ELLIOTT, AND A. KOH (2023): “Capability accumulation and conglomeratization in the information age,” *Journal of Economic Theory*, 210, 105647.
- COSTINOT, A. AND I. WERNING (2023): “Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation,” *The Review of Economic Studies*, 90, 2261–2291.
- CULLEN, Z., D. LI, AND S. LI (2025): “Labor as Capital: AI and the Ownership of Expertise,” *Working Paper*.
- DEMIRER, M., J. HORTON, N. IMMORLICA, B. LUCIER, AND P. SHAHIDI (2025): “The Economic Impacts of Generative AI on the Structure of Work,” Tech. rep.
- DWIVEDI, K. AND G. ROIG (2019): “Representation similarity analysis for efficient

- task taxonomy & transfer learning,” in *Proceedings of the IEEE / CVF conference on computer vision and pattern recognition*, 12387–12396.
- EROL, S., F. PARISE, AND A. TEYTELBOYM (2023): “Contagion in graphons,” *Journal of Economic Theory*, 211, 105673.
- FARBOODI, M. AND L. VELDKAMP (2021): “A model of the data economy,” Tech. rep., National Bureau of Economic Research Cambridge, MA, USA.
- GOLUB, B. (2025): “Eigenvalues in microeconomics,” *arXiv preprint arXiv:2502.12309*.
- GUERREIRO, J., S. REBELO, AND P. TELES (2022): “Should robots be taxed?” *The Review of Economic Studies*, 89, 279–311.
- HADFIELD, G. AND A. KOH (2025): “An Economy of AI Agents,” *NBER Handbook on the Economics of Transformative AI*, Forthcoming.
- HENDRYCKS, D., C. BURNS, S. BASART, A. ZOU, M. MAZEIKA, D. SONG, AND J. STEINHARDT (2020): “Measuring massive multitask language understanding,” *arXiv preprint arXiv:2009.03300*.
- HICKS, J. (1932): *The theory of wages*, Macmillan.
- HO, A., T. BESIROGLU, E. ERDIL, D. OWEN, R. RAHMAN, Z. C. GUO, D. ATKINSON, N. THOMPSON, AND J. SEVILLA (2024): “Algorithmic progress in language models,” *Advances in Neural Information Processing Systems*, 37, 58245–58283.
- HOFFMANN, J., S. BORGEAUD, A. MENSCH, E. BUCHATSKAYA, T. CAI, E. RUTHERFORD, D. D. L. CASAS, L. A. HENDRICKS, J. WELBL, A. CLARK, ET AL. (2022): “Training compute-optimal large language models,” *arXiv preprint arXiv:2203.15556*.
- JIANG, L. Y., X. C. LIU, N. P. NEJATIAN, M. NASIR-MOIN, D. WANG, A. ABIDIN, K. EATON, H. A. RIINA, I. LAUFER, P. PUNJABI, ET AL. (2023): “Health system-scale language models are all-purpose prediction engines,” *Nature*, 619, 357–362.
- JONES, B. F. AND X. LIU (2024): “A framework for economic growth with capital-embodied technical change,” *American Economic Review*, 114, 1448–1487.
- JONES, C. I. AND C. TONETTI (2020): “Nonrivalry and the Economics of Data,” *American Economic Review*, 110, 2819–2858.
- (2025): “Past Automation and Future AI: How Weak Links Tame the Growth Explosion,” Tech. rep., Working paper, Stanford GSB and NBER.
- KARPATY, A. (2025): “Tweet thread,” .
- KIM, J. W., T. Z. ZHAO, S. SCHMIDGALL, A. DEGUET, M. KOBILAROV, C. FINN, AND A. KRIEGER (2024): “Surgical robot transformer (srt): Imitation learning for surgical tasks,” *arXiv preprint arXiv:2407.12998*.
- KIM, K., S. KOTHA, P. LIANG, AND T. HASHIMOTO (2025): “Pre-training under infinite compute,” .
- KOH, A. AND S. MORRIS (2022): “Speed vs resilience in contagion,” *Available at SSRN 4292559*.
- KORINEK, A. AND D. SUH (2024): “Scenarios for the Transition to AGI,” Tech. rep.,

- National Bureau of Economic Research.
- KWA, T., B. WEST, J. BECKER, A. DENG, K. GARCIA, M. HASIN, S. JAWHAR, M. KINNIMENT, N. RUSH, S. VON ARX, R. BLOOM, T. BROADLEY, H. DU, B. GOODRICH, N. JURKOVIC, L. H. MILES, S. NIX, T. LIN, N. PARIKH, D. REIN, L. J. K. SATO, H. WIJK, D. M. ZIEGLER, E. BARNES, AND L. CHAN (2025): “Measuring AI Ability to Complete Long Tasks,” *arXiv preprint arXiv:2503.14499*, version 2.
- MORRIS, S. (2000): “Contagion,” *The Review of Economic Studies*, 67, 57–78.
- NARAYANAN, A. AND S. KAPOOR (2025): “AI as normal technology,” *Knight First Amend. Inst.*
- NORDHAUS, W. D. (2008): “Baumol’s diseases: a macroeconomic perspective,” *The BE Journal of Macroeconomics*, 8.
- (2021): “Are we approaching an economic singularity? information technology and the future of economic growth,” *American Economic Journal: Macroeconomics*, 13, 299–332.
- NYT (2025): “This A.I. Company Wants to Take Your Job,” .
- ORD, T. (2025): “The Extreme Inefficiency of RL for Frontier Models,” .
- PARISE, F. AND A. OZDAGLAR (2023): “Graphon games: A statistical framework for network games and interventions,” *Econometrica*, 91, 191–225.
- PATEL, D. (2025a): “RL is even more information inefficient than you thought,” .
- (2025b): *The Scaling Era: An Oral History of AI, 2019–2025*, Stripe Press.
- PAUL, K. AND J. HORWITZ (2026): “Meta to start capturing employee mouse movements, keystrokes for AI training data,” *Reuters*, accessed April 27, 2026.
- RESTREPO, P. (2025): “We Won’t be Missed: Work and Growth in the AGI World,” Tech. rep., National Bureau of Economic Research.
- ROMER, P. M. (1986): “Increasing returns and long-run growth,” *Journal of political economy*, 94, 1002–1037.
- SEVILLA, J. AND E. ROLDÁN (2024): “Training compute of frontier AI models grows by 4-5x per year,” .
- SMITH, H. L. (1995): *Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems: an introduction to the theory of competitive and cooperative systems*, 41, American Mathematical Soc.
- SOLOW, R. M. (1957): “Technical change and the aggregate production function,” *The review of Economics and Statistics*, 39, 312–320.
- TAO, T. (2006): *Nonlinear dispersive equations: local and global analysis*, 106, American Mathematical Soc.
- THE INFORMATION (2025): “How Anthropic and OpenAI Are Developing AI ‘Co-Workers’,” .
- TRAMMELL, P. (2025): “Workflows and Automation,” Tech. rep., Working Paper. Digital Economy Lab, Stanford University. [https . . .](https://...)
- TRAMMELL, P. AND A. KORINEK (2025): “Economic growth under transformative

- AI,” Tech. rep.
- U.S. SENATE COMMITTEE (2025): “The Big Tech Oligarchs’ War Against Workers: AI and Automation Could Destroy Nearly 100 Million U.S. Jobs in a Decade,” .
- WALL STREET JOURNAL (2025): “The AI Startup Fueling ChatGPT’s Expertise Is Now Valued at \$10 Billion,” .
- WIEDEMER, T., Y. LI, P. VICOL, S. S. GU, N. MATARESE, K. SWERSKY, B. KIM, P. JAINI, AND R. GEIRHOS (2025): “Video models are zero-shot learners and reasoners,” *arXiv preprint arXiv:2509.20328*.
- WILSON, R. (1975): “Informational economies of scale,” *The Bell Journal of Economics*, 184–195.
- WU, S., O. IRSOY, S. LU, V. DABRAVOLSKI, M. DREDZE, S. GEHRMANN, P. KAMBADUR, D. ROSENBERG, AND G. MANN (2023): “Bloomberggpt: A large language model for finance,” *arXiv preprint arXiv:2303.17564*.
- YANG, H., Y. ZHANG, J. XU, H. LU, P. A. HENG, AND W. LAM (2024): “Unveiling the generalization power of fine-tuned large language models,” *arXiv preprint arXiv:2403.09162*.
- ZEIRA, J. (1998): “Workers, machines, and economic growth,” *The Quarterly Journal of Economics*, 113, 1091–1117.

APPENDIX TO DATA-DRIVEN AUTOMATION

Organization. The appendix is organized as follows. [Appendix A](#) collects proofs of results in the main text. [Appendix B](#) provides a more detailed discussion of theoretical extensions alluded in the conclusion. [Appendix C](#) presents the details of the solution method behind numerical simulations. [Appendix D](#) collects additional numerical results.

A Proofs

A.1 [Section 3](#): Data Autarky

Proof of [Lemma 1](#). The final producer's first-order condition is $Y^{1/\sigma} \cdot y_i^{-1/\sigma} = p_i$, as usual.

To derive the equilibrium objects in each period t , notice that from the final producer's FOC, labor must be equally distributed across the non-automated tasks $(\gamma_t, 1]$, so labor allocation is determined by [Equations \(7\)](#) in the lemma.

The market clearing condition for capital can be written as

$$K_i \int_0^{\gamma_t} \frac{K_j}{K_i} dj = K$$

For any two automated i and j the final good producer's FOC implies

$$\frac{\psi_i^K K_i}{\psi_j^K K_j} = \left(\frac{r}{\psi_i^K} \right)^{-\sigma} \Rightarrow \frac{K_i}{K_j} = \left(\frac{\psi_i^K}{\psi_j^K} \right)^{\sigma-1}$$

so capital allocation is therefore given by [Equations \(6\)](#) in the lemma.

Next, plugging labor and capital allocations into the final producer's FOC and using the fact that intermediate tasks are perfectly competitive, we have that factor prices are thus given by [Equations \(8\)](#) and [\(9\)](#) in the lemma.

The automation boundary γ_t is then pinned down by a task that is indifferent between using labor and capital to produce i.e., $\psi_{\gamma_t}^K / r = \psi^L / L$ which can be implicitly rewritten as [Equation \(5\)](#) in the lemma. We will later verify the above expression admits a unique solution.

It remains to show regularity. From our data LOM,

$$D_{it} = D_{i0} + \int_0^t Y_{is} ds \quad \forall i \quad (14)$$

Since data is initially continuous and strictly decreasing in task index, by [Lemma 1](#), we have $\dot{D}_{it} = Y_{i0} \geq Y_{j0} = \dot{D}_{jt} \quad \forall i \leq j$. This then implies that in a small neighborhood around $t = 0$ the initial hypothesis that task data is weakly decreasing in task index survives, allowing us to conclude via the "Abstract Bootstrap Principle" ([Tao, 2006](#)). $\square$

Proof of Lemma 2. We will derive a generalized expression capturing arbitrary spillovers and show it reduces to the expression in the main text when data is task-specific. Rearranging (5), we have the following expression for capital productivity of the threshold task on the automation boundary:

$$\psi_{\gamma_t}^K = \frac{\psi^L L}{K} \frac{1}{1 - \gamma_t} \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \quad (15)$$

Suppose we run this system forward with $\gamma_t = \gamma$ held fixed. If the change in RHS weakly exceeds the change LHS, γ_t must be weakly decreasing in time given [Lemma 1](#) since a larger γ_t would make the RHS even larger and the LHS even smaller. Similarly, if the change in LHS strictly exceeds the change in RHS, γ_t must be strictly increasing in time to restore the equality. The partial of LHS w.r.t. time is

$$\partial_{\mathcal{A}_\gamma} \psi_\gamma^K \cdot \partial_t \mathcal{A}_\gamma(\mathbf{D}_t) \quad (16)$$

Where we abuse notation by using $\partial_{\mathcal{A}_\gamma}$ as the partial derivative w.r.t. effective data. Note that we write γ instead of γ_t as we are holding it fixed. The partial of RHS w.r.t. time is

$$\frac{\psi^L L}{K} \frac{\sigma - 1}{1 - \gamma} \int_0^\gamma \left(\frac{\psi_j^K}{\psi_\gamma^K} \right)^{\sigma-2} \frac{\psi_\gamma^K \left(\partial_{\mathcal{A}_j} \psi_j^K \right) \left(\partial_t \mathcal{A}_j(\mathbf{D}_t) \right) - \psi_j^K \left(\partial_{\mathcal{A}_\gamma} \psi_\gamma^K \right) \left(\partial_t \mathcal{A}_\gamma(\mathbf{D}_t) \right)}{\left(\psi_\gamma^K \right)^2} dj \quad (17)$$

Rearranging and substituting in (15), γ_t is weakly decreasing in time if and only if:

$$(\sigma - 1) \int_0^\gamma \left(\frac{\psi_j^K}{\psi_\gamma^K} \right)^{\sigma-1} \frac{\partial_t \log \mathcal{A}_j(\mathbf{D}_t)}{\partial_t \log \mathcal{A}_\gamma(\mathbf{D}_t)} dj \geq \sigma \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_\gamma^K} \right)^{\sigma-1} dj \quad (18)$$

Observe that in the case of no data spillovers, the expression reduces to

$$(\sigma - 1) \int_0^{\gamma_t} \left(\frac{f(j)}{f(\gamma_t)} \right)^{1/\eta} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{2\sigma-1-1/\eta} dj \geq \sigma \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \quad (19)$$

□

Proof of Proposition 1. First, all conditions for Lemma 1, Lemma 1, and Lemma 2 are met, so equilibrium characterizations from those results apply. By Lemma 1, we know γ_t is monotonically increasing since $\sigma \leq 1$. We will first show that $\lim_{t \rightarrow \infty} \gamma_t = 1$ when $\sigma < 1/\eta$ for any $W \in \mathcal{W}$ before establishing that the ratio of data is a positive constant for any two tasks in the limit.

Full limit automation (gross complements). We can rewrite (5) as

$$\psi_{\gamma_t}^K = \frac{\psi^L L}{K} \frac{1}{1 - \gamma_t} \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \leq \frac{\psi^L L}{K} \frac{\gamma_t}{1 - \gamma_t} \quad (20)$$

Here the inequality follows from the fact that $\psi_j^K \geq \psi_{\gamma_t}^K$ and $\sigma \leq 1$. Since $\psi_{\gamma_t}^K \geq \psi_1^K$ by Lemma 1, it suffices to show that

$$\lim_{t \rightarrow \infty} \psi_1^K(\mathcal{A}_1(\mathbf{D}_t)) = \infty \quad (21)$$

This is indeed the case since all tasks have the outside option of producing with labor, which yields a task quantity of

$$\psi_L \frac{L}{1 - \gamma_t} \geq \psi^L \frac{L}{1 - \gamma_0} > 0 \quad (22)$$

From our data LOM it is immediate that effective data for task 1—and therefore all tasks—go to infinity, so we must approach full automation in the limit.

Full limit automation (gross substitutes with $\sigma \leq 1/\eta$). Again, we know $\psi_{\gamma_t}^K$ can be written equivalently as

$$\psi_{\gamma_t}^K = \frac{\psi^L L}{K} \frac{1}{1 - \gamma_t} \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \quad (23)$$

The same argument from the gross complements case shows that the LHS goes to infinity as $t \rightarrow \infty$. Towards a contradiction, suppose γ_t is bounded above by some

$U < 1$,

$$\frac{\psi^L L}{K} \frac{1}{1 - \gamma_t} \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \leq \frac{\psi^L L}{K} \frac{1}{1 - U} \left(\frac{\psi_0^K}{\psi_1^K} \right)^{\sigma-1} \quad (24)$$

Here the inequality uses $\sigma \geq 1$ and the monotonicity lemma. It suffices to now show ψ_0^K/ψ_1^K is bounded for a contradiction. Differentiating the ratio of productivities w.r.t. time,

$$\begin{aligned} \frac{d \ln(\psi_0^K/\psi_1^K)}{dt} &= \frac{(\psi_0^K)'}{\psi_0^K} \int_0^1 W(0, j) Y_j dj - \frac{(\psi_1^K)'}{\psi_1^K} \int_0^1 W(1, j) Y_j dj \\ &= \eta \left(\frac{f(0)}{\psi_0^K} \right)^{\frac{1}{\eta}} \int_0^1 W(0, j) Y_j dj - \eta \left(\frac{f(1)}{\psi_1^K} \right)^{\frac{1}{\eta}} \int_0^1 W(1, j) Y_j dj \\ &= \eta \left(\frac{f(1)}{\psi_1^K} \right)^{\frac{1}{\eta}} \left(\int_0^1 W(1, j) Y_j dj \right) \left(\left(\frac{\psi_1^K f(0)}{\psi_0^K f(1)} \right)^{\frac{1}{\eta}} \frac{\int_0^1 W(0, j) Y_j dj}{\int_0^1 W(1, j) Y_j dj} - 1 \right) \end{aligned} \quad (25)$$

Here we write ψ_i^K for brevity, suppressing the time subscript. Note the ratio decreases wherever

$$\left(\frac{\psi_{1,t}^K f(0)}{\psi_{0,t}^K f(1)} \right)^{\frac{1}{\eta}} \frac{\int_0^1 W(0, k) Y_{k,t} dk}{\int_0^1 W(1, k) Y_{k,t} dk} \leq 1$$

Note that

$$\frac{\int_0^1 W(0, k) Y_{k,t} dk}{\int_0^1 W(1, k) Y_{k,t} dk} \leq \frac{Y_{0,t} \int_0^1 W(0, k) dk}{Y_{1,t} \int_0^1 W(1, k) dk} \quad (26)$$

$$\leq \left(\frac{\psi_0^K}{\psi_1^K} \right)^{\sigma} \frac{\int_0^1 W(0, k) dk}{\int_0^1 W(1, k) dk} \quad \text{by Lemma A.1} \quad (27)$$

All supporting lemmas are collected at the end of proofs to propositions for which they are first introduced. Applying the upperbound from above, the ratio will also decrease when the following condition holds

$$\left(\frac{f(0)}{f(1)} \right)^{\frac{1}{\eta}} \left(\frac{\psi_{0,t}^K}{\psi_{1,t}^K} \right)^{\frac{\sigma\eta-1}{\eta}} \frac{\int_0^1 W(0, k) dk}{\int_0^1 W(1, k) dk} \leq 1 \iff \frac{\psi_{0,t}^K}{\psi_{1,t}^K} \geq \left(\frac{\int_0^1 W(0, k) dk}{\int_0^1 W(1, k) dk} \left(\frac{f(0)}{f(1)} \right)^{\frac{1}{\eta}} \right)^{\frac{\eta}{1-\sigma\eta}} \quad (28)$$

The RHS is bounded by assumption, so we conclude that

$$\frac{\psi_0^K(\mathbf{D}_{0,t})}{\psi_1^K(\mathbf{D}_{1,t})} \leq \max \left\{ \frac{\psi_0^K(\mathbf{D}_{0,0})}{\psi_1^K(\mathbf{D}_{1,0})}, \left(\frac{\int_0^1 W(0,k) dk}{\int_0^1 W(1,k) dk} \left(\frac{f(0)}{f(1)} \right)^{\frac{1}{\eta}} \right)^{\frac{\eta}{1-\sigma\eta}} \right\} < \infty \quad \forall t \geq 0 \quad (29)$$

This concludes the proof for the full limit automation result. We now turn to establishing that data will be balanced in the limit.

Balanced data. We now establish that the relative data ratio between any two tasks $i \neq j$ will be balanced.

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{D_{it}}{D_{jt}} &= \lim_{t \rightarrow \infty} \frac{y_{it}}{y_{jt}} && \text{(L'Hopital's)} \\ &= \lim_{t \rightarrow \infty} \frac{\psi_{it}^K K_{it}}{\psi_{jt}^K K_{jt}} && \text{(tasks will be produced by capital asymptotically)} \\ &= \lim_{t \rightarrow \infty} \left(\frac{\psi_{it}^K}{\psi_{jt}^K} \right)^\sigma && (K_{it} = \frac{(\psi_{it}^K)^{\sigma-1}}{\int_0^\gamma (\psi_j^K)^{\sigma-1} dj} K) \\ &= \lim_{t \rightarrow \infty} \left(\frac{f_i}{f_j} \right)^\sigma \left(\frac{D_{it}}{D_{jt}} \right)^{\eta\sigma} \end{aligned}$$

Rearranging gives the expression for $\lim_{t \rightarrow \infty} D_{it}/D_{jt}$. Observe that with or without capital accumulation, we will have $r/w \rightarrow \infty$. Note that this argument goes through with capital accumulation, since we have full automation so long as capital is bounded away from 0 (See [Equation \(5\)](#)).

Lemma A.1 (An upperbound on the ratio of task production).

$$\frac{Y_0}{Y_j} \leq \left(\frac{\psi_0^K}{\psi_j^K} \right)^\sigma \quad (30)$$

Proof of Lemma A.1. We know the most productive task will always be produced with capital, so

$$Y_0 = \frac{(\psi_0^K)^\sigma}{\int_0^\gamma (\psi_j^K)^{\sigma-1} dj} K \quad (31)$$

In addition, if task j is produced with capital, the above expression holds with 0

replaced by j . Otherwise, we know

$$Y_j = Y_\gamma = \frac{(\psi_\gamma^K)^\sigma}{\int_0^\gamma (\psi_j^K)^{\sigma-1} dj} K \geq \frac{(\psi_j^K)^\sigma}{\int_0^\gamma (\psi_j^K)^{\sigma-1} dj} K \quad (32)$$

Dividing the two expressions, we have

$$\frac{Y_0}{Y_j} \leq \left(\frac{\psi_0^K}{\psi_j^K} \right)^\sigma \quad (33)$$

as desired. □

□

Proof of Proposition 2. First, all conditions for [Lemma 1](#), [Lemma 1](#), and [Lemma 2](#) are met, so equilibrium characterizations from those results apply. We will first show the upperbound result before establishing that data will be imbalanced in the limit.

Bounded automation. We will prove

$$\gamma_t \leq \max\{\gamma_0, L_{f,\sigma}\} < 1 \quad \forall t \geq 0 \quad (34)$$

We do so by establishing that $\gamma_t \geq L_{f,\sigma}$ implies γ_t is weakly decreasing in time. We remind the reader $L_{f,\sigma}$ comes from the σ -regularity condition, which stipulates that there exists some constant $L_{\sigma,f} \in (0, 1)$ such that

$$\frac{\sigma}{\sigma-1} f(i) < \frac{1}{i} \int_0^i f(j) dj \quad \text{for all } i \geq L_{\sigma,f}. \quad (35)$$

From [Lemma 2](#), we know γ_t is weakly decreasing in time if

$$(\sigma-1) \int_0^{\gamma_t} \left(\frac{f(j)}{f(\gamma_t)} \right)^{1/\eta} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{2\sigma-1-1/\eta} dj \geq \sigma \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \quad (36)$$

Observe that

$$\int_0^{\gamma_t} f(j) \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \geq \frac{1}{\gamma_t} \left(\int_0^{\gamma_t} f(j) dj \right) \left(\int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \right) \quad (37)$$

$$\geq \frac{\sigma}{\sigma-1} f(\gamma_t) \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj \quad (38)$$

where the first inequality follows from Chebyshev/FKG (see, e.g., Theorem 1 in [Armstrong \(1993/94\)](#)); the second is due to f being σ -regular and $\gamma_t \geq L_{f,\sigma}$ by assumption. Dividing both sides of the expression above by $f(\gamma_t)$ concludes the proof since $f(j) \geq f(\gamma_t) > 0 \forall j \leq \gamma_t$ and $1/\eta > 1$.

Imbalanced data. It is easy to verify D_{it}/D_{jt} is monotonically increasing over time, so if D_{it}/D_{jt} does not diverge, the limit exists by monotone convergence. First, if j is always produced using capital after some time T , then following the same steps as before

$$\lim_{t \rightarrow \infty} \frac{D_{0t}}{D_{jt}} = \left(\frac{f_0}{f_j} \right)^\sigma \lim_{t \rightarrow \infty} \left(\frac{D_{0t}}{D_{jt}} \right)^{\eta\sigma}$$

Suppose the ratio is bounded, then the limit exists, but the RHS is strictly larger than the LHS. Moreover, we know $D_{0t}/D_{jt} \geq 1$, so the inequality yields a contradiction. Therefore the limit must diverge to infinity in this case. On the other hand, if j is produced only using labor only after some time T , then

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{D_{0,t}}{D_{jt}} &= \lim_{t \rightarrow \infty} \frac{Y_{0,t}}{Y_{jt}} \\ &\geq \lim_{t \rightarrow \infty} \frac{Y_{0,t}}{\psi^{L/L/(1-j)}} \quad (\text{all tasks above } j \text{ must be produced using labor}) \\ &\geq \lim_{t \rightarrow \infty} \frac{1}{\psi^{L/L/(1-j)}} \underbrace{\left(\frac{\left(\psi_{0,t}^K \right)^{\sigma-1}}{\int_0^{\gamma_t} \left(\psi_{k,t}^K \right)^{\sigma-1} dk} \right)}_{\geq 1} \psi_{0,t}^K \\ &\geq \lim_{t \rightarrow \infty} \frac{1}{\psi^{L/L/(1-j)}} \psi_{0,t}^K = \infty \end{aligned}$$

This concludes the proof. □

Proof of Proposition 3. Recall wage is given by

$$w_t = Y_t^{\frac{1}{\sigma}} \left(\frac{(1 - \gamma_t) (\psi^L)^{\sigma-1}}{L} \right)^{\frac{1}{\sigma}}$$

so the long-run wage is governed by $Y_t(1 - \gamma_t)$ as the rest are constants. Using our expression for output, we have

$$Y_t(1 - \gamma_t) = \left(\gamma_t^{\frac{1}{\sigma}} ((1 - \gamma_t) \Psi_K(\gamma_t, \boldsymbol{\psi}_t) K_t)^{\frac{\sigma-1}{\sigma}} + (1 - \gamma_t) (\psi_L L)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}.$$

First consider the case where $\sigma \leq 1/\eta$. Note that the second term is going to zero in the limit. As for the first term, we need to evaluate $(1 - \gamma_t) \Psi_K(\gamma_t, \boldsymbol{\psi}_t) K_t$. Using our implicit definition of γ_t , we can write

$$1 - \gamma_t = \frac{1}{\psi_{\gamma_t}^K} \frac{\psi^L L}{K_t} \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj$$

So

$$(1 - \gamma_t) \Psi_K(\gamma_t, \boldsymbol{\psi}_t) K_t = \psi^L L \frac{\Psi_K(\gamma_t, \boldsymbol{\psi}_t)}{\psi_{\gamma_t}^K} \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj$$

When $\sigma \leq 1/\eta$, we have balanced data, so

$$\lim_{t \rightarrow \infty} \frac{\Psi_K(\gamma_t, \boldsymbol{\psi}_t)}{\psi_{\gamma_t}^K} \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj = C < \infty$$

On the other hand, when $\sigma > 1/\eta$, we know $\lim_{t \rightarrow \infty} Y_t = \infty$ and γ_t will be upper-bounded per [Proposition 2](#), so the long-run wage therefore diverges to infinity. Note that this argument only requires the balanced data result from [Proposition 1](#), which holds with or without capital accumulation as noted in the proof. □

Proof of Proposition 4. Under our maintained assumptions on the boundedness of f and D_0 , there are constants $\underline{f}, \bar{f}, \underline{D}_0, \bar{D}_0 > 0$ such that

$$\underline{f} \leq f_i \leq \bar{f}, \quad \underline{D}_0 \leq D_{i0} \leq \bar{D}_0 \quad \forall i \in X.$$

We first argue that this delivers a whole-path bound. Fix $i < j$ and write $x_{ij,t} :=$

D_{it}/D_{jt} . The balanced-data ratio is

$$x_{ij}^* := \left(\frac{f_i}{f_j} \right)^{\frac{\sigma}{1-\sigma\eta}}.$$

Suppose $1 \leq x_{ij,0} \leq x_{ij}^*$. We claim that

$$1 \leq x_{ij,t} \leq x_{ij}^* \quad \forall t \geq 0.$$

To see this, observe that $x_{ij,t}$ can cross the upper boundary only if $\dot{x}_{ij,t} > 0$ when $x_{ij,t} = x_{ij}^*$. If both tasks are automated, then

$$\frac{y_{it}}{y_{jt}} = \left(\frac{\psi_i^K}{\psi_j^K} \right)^\sigma = \left(\frac{f_i}{f_j} \right)^\sigma x_{ij,t}^{\eta\sigma}.$$

Hence

$$\frac{\dot{x}_{ij,t}}{x_{ij,t}} = \frac{y_{it}}{D_{it}} - \frac{y_{jt}}{D_{jt}} = \frac{y_{jt}}{D_{jt}} \left[\left(\frac{f_i}{f_j} \right)^\sigma x_{ij,t}^{\eta\sigma-1} - 1 \right],$$

which is zero at $x_{ij,t} = x_{ij}^*$ and negative above it because $\eta\sigma < 1$. If both tasks are produced by labor, then $y_{it} = y_{jt}$, so the ratio moves toward one. Finally, if i is automated and j is produced by labor, then $y_{it}/y_{jt} \leq (\psi_i^K/\psi_j^K)^\sigma$, so the same upper-bound argument applies. Thus $x_{ij,t}$ cannot cross above x_{ij}^* . A symmetric argument at $x_{ij,t} = 1$, using $f_i \geq f_j$ and the monotonicity of capital productivity, shows it cannot cross below one.

It follows that, for every $i < j$ and every $t \geq 0$,

$$\frac{\psi_i^K}{\psi_j^K} = \frac{f_i}{f_j} x_{ij,t}^\eta \leq \left(\frac{f_i}{f_j} \right)^{\frac{1}{1-\sigma\eta}}.$$

Since f_i is bounded above and away from zero, the right-hand side is uniformly bounded. In particular, there exists $R < \infty$ such that

$$1 \leq \frac{\psi_j^K}{\psi_{\gamma_t}^K} \leq R \quad \forall j \leq \gamma_t, t \geq 0.$$

Under this stronger initial-condition restriction, we can take $T_R = 0$ in the argument below. Since the automation boundary is weakly increasing along the path, we can also take the lower bound on γ_t to be γ_0 , and therefore take $T = 0$. The rest of the proof is written for a generic date T . In the non-asymptotic case just described, the same argument applies with $T = 0$, $\gamma_\star = \gamma_0$, $\underline{B} = \underline{D}_0$, and $\overline{B} = \overline{D}_0$.

Since $\sigma < 1/\eta$, **Proposition 1** implies full limit automation and balanced data. Since f_i is bounded above and away from zero, the bounded relative-productivity conclusion in **Proposition 1** gives an $R < \infty$ and a finite time T_R such that for every $t \geq T_R$ and every automated task $j \leq \gamma_t$,

$$1 \leq \frac{\psi_j^K}{\psi_{\gamma_t}^K} \leq R.$$

The lower bound uses monotonicity of capital productivity in the task index. Fix any $\gamma_\star \in (0, 1)$. Full limit automation gives a finite time $T_\gamma(\gamma_\star)$ such that $\gamma_t \geq \gamma_\star$ for all $t \geq T_\gamma(\gamma_\star)$. Take $T = \max\{T_R, T_\gamma(\gamma_\star)\}$. The bounds below are non-asymptotic for every $t \geq T$, but the date T is determined by the eventual convergence results in **Proposition 1**. Thus $\gamma_\star$ is arbitrary, but fixing a larger $\gamma_\star$ may require a later $T_\gamma(\gamma_\star)$.

Define

$$M_t := \int_0^{\gamma_t} \left(\frac{\psi_j^K}{\psi_{\gamma_t}^K} \right)^{\sigma-1} dj.$$

The bounded-ratio condition implies the stated bounds on M_t . If $\sigma \geq 1$, the integrand in M_t lies between 1 and $R^{\sigma-1}$, so

$$\gamma_\star \leq \gamma_t \leq M_t \leq \gamma_t R^{\sigma-1} \leq R^{\sigma-1}.$$

If $\sigma < 1$, the integrand lies between $R^{\sigma-1}$ and 1, so

$$\gamma_\star R^{\sigma-1} \leq \gamma_t R^{\sigma-1} \leq M_t \leq \gamma_t \leq 1.$$

Thus $0 < \underline{M} \leq M_t \leq \overline{M} < \infty$.

By the static equilibrium characterization, the automation boundary satisfies

$$\psi_{\gamma_t}^K = \left(\frac{L}{K(1-\gamma_t)} \int_0^{\gamma_t} (\psi_j^K)^{\sigma-1} dj \right)^{1/\sigma}.$$

Equivalently,

$$1 - \gamma_t = \frac{L}{K} \frac{M_t}{\psi_{\gamma_t}^K}.$$

Since the automation boundary is weakly increasing, the marginal task at time t has accumulated data from labor production until it reaches the automation boundary. Define cumulative labor-produced data for not-yet-automated tasks by

$$H_t := \int_0^t \frac{L}{1-\gamma_s} ds.$$

Then

$$D_{\gamma_t, t} = D_{\gamma_t, 0} + H_t,$$

and the boundedness of the initial data stock implies

$$\underline{B}_t := \underline{D}_0 + H_t \leq D_{\gamma_t, t} \leq \overline{D}_0 + H_t =: \overline{B}_t.$$

Let

$$\underline{B} := \underline{B}_T, \quad \overline{B} := \overline{B}_T.$$

Therefore

$$\psi_{\gamma_t}^K = f_{\gamma_t} D_{\gamma_t, t}^\eta.$$

It follows that

$$1 - \gamma_t = \frac{L}{K} \frac{M_t}{f_{\gamma_t}} D_{\gamma_t, t}^{-\eta}.$$

Using $f_{\gamma_t} \in [\underline{f}, \overline{f}]$, the bounds on M_t , and the envelope for $D_{\gamma_t, t}$,

$$\frac{LM}{K\underline{f}} \overline{B}_t^{-\eta} \leq 1 - \gamma_t \leq \frac{L\overline{M}}{K\underline{f}} \underline{B}_t^{-\eta}.$$

Since $\dot{H}_t = L/(1 - \gamma_t)$, $\dot{\underline{B}}_t = \dot{\overline{B}}_t = \dot{H}_t$. The previous inequalities imply

$$\frac{Kf}{\overline{M}} \underline{B}_t^\eta \leq \dot{\underline{B}}_t = \dot{\overline{B}}_t \leq \frac{K\overline{f}}{\underline{M}} \overline{B}_t^\eta.$$

Compare $\underline{B}_t$ and $\overline{B}_t$ to the two ODEs

$$\dot{B}_t^- = \frac{Kf}{\overline{M}} (B_t^-)^\eta, \quad \dot{B}_t^+ = \frac{K\overline{f}}{\underline{M}} (B_t^+)^\eta, \quad B_T^- = \underline{B}, \quad B_T^+ = \overline{B}.$$

Because $0 < \eta < 1$, these comparison equations have closed-form solutions

$$B_t^- = \left(\underline{B}^{1-\eta} + (1-\eta) \frac{Kf}{\overline{M}} (t-T) \right)^{\frac{1}{1-\eta}},$$

and

$$B_t^+ = \left(\overline{B}^{1-\eta} + (1-\eta) \frac{K\overline{f}}{\underline{M}} (t-T) \right)^{\frac{1}{1-\eta}}.$$

The comparison theorem gives $B_t^- \leq \underline{B}_t$ and $\overline{B}_t \leq B_t^+$. Combining this with the previ-

ous bounds on $1 - \gamma_t$ yields

$$1 - \gamma_t \geq \frac{LM}{Kf} (B_t^+)^{-\eta}, \quad 1 - \gamma_t \leq \frac{LM}{Kf} (B_t^-)^{-\eta},$$

which is the desired pair of power-law bounds. Finally, consider the stronger initial-condition restriction in the statement. The invariant-region argument at the beginning of the proof gives a constant $R < \infty$ such that

$$1 \leq \frac{\psi_j^K}{\psi_{\gamma_t}^K} \leq R \quad \forall j \leq \gamma_t, t \geq 0.$$

Thus no asymptotic appeal to **Proposition 1** is needed to obtain the relative-productivity bound. Since γ_t is weakly increasing, $\gamma_t \geq \gamma_0$ for all $t \geq 0$, and therefore

$$0 < \underline{M} \leq M_t \leq \overline{M} < \infty \quad \forall t \geq 0,$$

where one may take

$$(\underline{M}, \overline{M}) = \begin{cases} (\gamma_0, R^{\sigma-1}), & \sigma \geq 1, \\ (\gamma_0 R^{\sigma-1}, 1), & \sigma < 1. \end{cases}$$

Moreover $H_0 = 0$, so the initial envelopes are simply $\underline{B}_0 = \underline{D}_0$ and $\overline{B}_0 = \overline{D}_0$. Repeating the same comparison argument above with $T = 0$, $\underline{B} = \underline{D}_0$, and $\overline{B} = \overline{D}_0$ gives the displayed bounds for every $t \geq 0$. $\square$

A.2 Section 4: Contagious Automation

Proof of Proposition 5. The argument is to show that connected spillovers eventually prevent any task from remaining permanently data-poor. We first show that aggregate raw data grows without bound. We then use uniform strong connectedness to propagate a lower bound on effective data through the graphon in finitely many doubling intervals. This eventually makes every task's effective data comparable to aggregate raw data. Hence all capital productivities diverge and remain uniformly comparable. The final step rules out a positive-measure set of labor-produced tasks, because capital eventually becomes strictly cheaper than labor on almost every task.

For $S \subseteq X$ and $q > 0$, define

$$\Gamma_q(S) := \left\{ i \in X : \int_S W(i, j) dj \geq q \right\}.$$

Step 1: Showing aggregate raw data diverges. Let

$$M_t := \int_X D_{i,t} di$$

denote aggregate raw data. A labor-only allocation with $L_i = L$ for all i is feasible and delivers final output $\psi^L L$. Thus equilibrium output is bounded below by $\psi^L L$. Since the CES aggregate is no larger than $\int_X y_{i,t} di$ on the unit-measure task space,

$$\dot{M}_t = \int_X y_{i,t} di \geq \psi^L L.$$

Therefore $M_t \rightarrow \infty$.

Step 2: Propagating effective data through the network. We first prove a simple implication: if task i has effective data at least a fixed fraction of aggregate data, then task i 's output is at least a fixed fraction of aggregate raw-data growth.

Fix $b \in (0, 1]$, and define

$$\lambda(b) := \frac{f}{\underline{f}} b^\eta.$$

Since $b \leq 1$, we have $\lambda(b) \leq 1$. Suppose

$$\mathcal{A}_i(\mathbf{D}_t) \geq b M_t.$$

Because $W \leq 1$,

$$\mathcal{A}_j(\mathbf{D}_t) \leq M_t \quad \text{for a.e. } j.$$

Thus, using $\psi_{j,t}^K = f_j \mathcal{A}_j(\mathbf{D}_t)^\eta$,

$$\psi_{i,t}^K \geq \underline{f} b^\eta M_t^\eta = \lambda(b) \bar{f} M_t^\eta \geq \lambda(b) \psi_{j,t}^K \quad \text{for a.e. } j.$$

This is a comparison of capital productivities as technologies; it does not assume that either task is actually produced by capital. Let

$$\chi_{j,t} := \min \left\{ \frac{r_t}{\psi_{j,t}^K}, \frac{w_t}{\psi_{j,t}^L} \right\}$$

be the unit cost of task j . The productivity comparison implies

$$\frac{r_t}{\psi_{i,t}^K} \leq \lambda(b)^{-1} \frac{r_t}{\psi_{j,t}^K}.$$

Since $\lambda(b) \leq 1$,

$$\chi_{i,t} \leq \lambda(b)^{-1} \chi_{j,t} \quad \text{for a.e. } j.$$

CES demand therefore gives

$$y_{i,t} \geq \lambda(b)^\sigma y_{j,t} \quad \text{for a.e. } j. \quad (39)$$

Integrating [Equation \(39\)](#) over j , we obtain

$$y_{i,t} \geq \lambda(b)^\sigma \dot{M}_t. \quad (40)$$

We now repeatedly use this one-step implication. It will be convenient to do a time change by using aggregate-data doubling time: over each interval on which M_t doubles, any task with effective data comparable to M_t produces a fixed fraction of the new aggregate raw data. This allows the lower bound to move from one spillover layer to the next. Fix T_0 with $M_{T_0} > 0$. Define the following doubling times

$$\tau_m := \inf\{t \geq T_0 : M_t \geq 2^m M_{T_0}\} \quad \text{for } m = 0, 1, 2, \dots$$

Then

$$M_{\tau_{m+1}} = 2M_{\tau_m}.$$

For each m , define

$$h_r^m(i) := \frac{1}{M_{\tau_m}} \int_X W^{(r)}(i, j) D_{j, \tau_m} dj, \quad r = 1, \dots, n.$$

Since $0 \leq W \leq 1$ and $\mu(X) = 1$, induction gives $0 \leq W^{(r)} \leq 1$ for every r . Hence

$$0 \leq h_r^m(i) \leq 1 \quad \text{for a.e. } i.$$

By the recursive definition of $W^{(r)}$ and Tonelli's theorem,

$$h_{r+1}^m(i) = \int_X W(i, k) h_r^m(k) dk. \quad (41)$$

Uniform strong connectedness gives $h_n^m(i) \geq \varepsilon$ for a.e. i . Set $a_n := \min\{\varepsilon, 1\}$, and recursively set

$$a_r := \frac{a_{r+1}}{2}, \quad r = n-1, \dots, 1.$$

Let

$$S_r^m := \{i \in X : h_r^m(i) \geq a_r\}.$$

Then $S_n^m = X$ up to a null set, and

$$\mathcal{A}_i(\mathbf{D}_{\tau_m}) \geq a_1 M_{\tau_m} \quad \text{for a.e. } i \in S_1^m.$$

We next show that if a task has high $(r+1)$ -step exposure, then a nontrivial amount of its direct spillover weight must fall on tasks with high r -step exposure:

$$S_{r+1}^m \subseteq \Gamma_{a_r}(S_r^m) \quad \text{up to null sets, for each } r < n. \quad (42)$$

Indeed, if $i \in S_{r+1}^m$, then by [Equation \(41\)](#),

$$a_{r+1} \leq h_{r+1}^m(i) = \int_X W(i, k) h_r^m(k) dk.$$

On $X \setminus S_r^m$, $h_r^m < a_r$, while on S_r^m , $h_r^m \leq 1$. Therefore

$$h_{r+1}^m(i) \leq a_r + \int_{S_r^m} W(i, k) dk.$$

Since $a_{r+1} = 2a_r$, it follows that

$$\int_{S_r^m} W(i, k) dk \geq a_r,$$

which proves [Equation \(42\)](#). The sets S_r^m identify where the spillover exposure is large. We now choose constants b_r so that membership in S_r^m implies an actual lower bound on effective data after $r-1$ doubling intervals. Define constants $b_1, \dots, b_n$ by

$$b_1 := a_1, \quad b_{r+1} := \frac{a_r}{2} \left[\lambda \left(\frac{b_r}{2} \right) \right]^\sigma, \quad r = 1, \dots, n-1.$$

By construction, $0 < b_r \leq 1$ for every r . We claim that for each $r = 1, \dots, n$,

$$\mathcal{A}_i(\mathbf{D}_{\tau_{m+r-1}}) \geq b_r M_{\tau_{m+r-1}} \quad \text{for a.e. } i \in S_r^m. \quad (43)$$

The case $r = 1$ follows from the definition of S_1^m and $b_1 = a_1$. Suppose [Equation \(43\)](#) holds for some $r < n$. For $u \in [\tau_{m+r-1}, \tau_{m+r}]$, data is nondecreasing and $M_u \leq 2M_{\tau_{m+r-1}}$. Hence, for a.e. $i \in S_r^m$,

$$\mathcal{A}_i(\mathbf{D}_u) \geq \frac{b_r}{2} M_u.$$

Applying [Equation \(40\)](#) with $b = b_r/2$, we get

$$y_{i,u} \geq \left[\lambda \left(\frac{b_r}{2} \right) \right]^\sigma M_u \quad \text{for a.e. } i \in S_r^m.$$

Integrating over $u \in [\tau_{m+r-1}, \tau_{m+r}]$ and dropping the nonnegative initial data term,

$$D_{i, \tau_{m+r}} \geq \left[\lambda \left(\frac{b_r}{2} \right) \right]^\sigma (M_{\tau_{m+r}} - M_{\tau_{m+r-1}}).$$

Since $M_{\tau_{m+r}} = 2M_{\tau_{m+r-1}}$,

$$D_{i, \tau_{m+r}} \geq \frac{1}{2} \left[\lambda \left(\frac{b_r}{2} \right) \right]^\sigma M_{\tau_{m+r}} \quad \text{for a.e. } i \in S_r^m. \quad (44)$$

Now take $i \in S_{r+1}^m$. By [Equation \(42\)](#),

$$\int_{S_r^m} W(i, k) dk \geq a_r.$$

Using [Equation \(44\)](#),

$$\begin{aligned} \mathcal{A}_i(\mathbf{D}_{\tau_{m+r}}) &= \int_X W(i, k) D_{k, \tau_{m+r}} dk \\ &\geq \int_{S_r^m} W(i, k) D_{k, \tau_{m+r}} dk \\ &\geq \frac{1}{2} \left[\lambda \left(\frac{b_r}{2} \right) \right]^\sigma M_{\tau_{m+r}} \int_{S_r^m} W(i, k) dk \\ &\geq \frac{a_r}{2} \left[\lambda \left(\frac{b_r}{2} \right) \right]^\sigma M_{\tau_{m+r}} \\ &= b_{r+1} M_{\tau_{m+r}}. \end{aligned}$$

This gives [Equation \(43\)](#).

Since $S_n^m = X$ up to a null set, [Equation \(43\)](#) with $r = n$ gives, for every m ,

$$\mathcal{A}_i(\mathbf{D}_{\tau_{m+n-1}}) \geq b_n M_{\tau_{m+n-1}} \quad \text{for a.e. } i \in X. \quad (45)$$

We now extend the bound from doubling times to all sufficiently large times. Take any $t \geq \tau_{n-1}$, and let $s \geq n-1$ be such that $t \in [\tau_s, \tau_{s+1}]$. Set

$$m = s - n + 1.$$

Then $m + n - 1 = s$, so [Equation \(45\)](#), applied with this starting m , gives

$$\mathcal{A}_i(\mathbf{D}_{\tau_s}) \geq b_n M_{\tau_s} \quad \text{for a.e. } i \in X.$$

Since data is nondecreasing,

$$\mathcal{A}_i(\mathbf{D}_t) \geq \mathcal{A}_i(\mathbf{D}_{\tau_s}).$$

And since $t \in [\tau_s, \tau_{s+1}]$,

$$M_t \leq M_{\tau_{s+1}} = 2M_{\tau_s}.$$

Therefore

$$\mathcal{A}_i(\mathbf{D}_t) \geq \frac{b_n}{2} M_t \quad \text{for a.e. } i \in X.$$

Since $W \leq 1$, also

$$\mathcal{A}_i(\mathbf{D}_t) \leq M_t \quad \text{for a.e. } i \in X.$$

Thus, for all $t \geq \tau_{n-1}$,

$$\frac{b_n}{2} M_t \leq \mathcal{A}_i(\mathbf{D}_t) \leq M_t \quad \text{for a.e. } i \in X. \quad (46)$$

Step 3: Comparing productivities. By [Equation \(46\)](#), every task's effective data is eventually bounded above and below by fixed multiples of M_t . Since $M_t \rightarrow \infty$, every task's capital productivity diverges, and cross-task capital productivities remain uniformly comparable.

Let

$$\underline{\psi}_t := \operatorname{ess\,inf}_{i \in X} \psi_{i,t}^K.$$

Then

$$\underline{\psi}_t \geq f \left(\frac{b_n}{2} M_t \right)^\eta \rightarrow \infty.$$

Moreover,

$$\operatorname{ess\,sup}_{i \in X} \psi_{i,t}^K \leq \bar{f} M_t^\eta \leq C_\psi \underline{\psi}_t$$

for some finite constant C_ψ .

Step 4: Ruling out a positive labor share. Suppose, toward a contradiction, that full automation fails. Then there exist $\delta > 0$ and dates $t_k \rightarrow \infty$ such that

$$\mu(\mathcal{L}_{t_k}) \geq \delta.$$

At such dates, CES demand gives $y_{i,t_k} = C_{t_k} c_{i,t_k}^{-\sigma}$, where c_{i,t_k} is the unit cost of task i . On $\mathcal{K}_{t_k}$, $c_{i,t_k} = r_{t_k} / \psi_{i,t_k}^K$, so capital market clearing gives

$$K = C_{t_k} r_{t_k}^{-\sigma} \int_{\mathcal{K}_{t_k}} (\psi_{i,t_k}^K)^{\sigma-1} di.$$

On $\mathcal{L}_{t_k}$, $c_{i,t_k} = w_{t_k}/\psi^L$, so labor market clearing gives

$$L = C_{t_k} w_{t_k}^{-\sigma} \mu(\mathcal{L}_{t_k}) (\psi^L)^{\sigma-1}.$$

Dividing the two market-clearing equations,

$$\frac{r_{t_k}}{w_{t_k}} = \left[\frac{L \int_{\mathcal{K}_{t_k}} (\psi_{i,t_k}^K)^{\sigma-1} di}{K \mu(\mathcal{L}_{t_k}) (\psi^L)^{\sigma-1}} \right]^{1/\sigma}.$$

By Step 3, for some finite C_σ ,

$$\int_{\mathcal{K}_{t_k}} (\psi_{i,t_k}^K)^{\sigma-1} di \leq C_\sigma \psi_{t_k}^{\sigma-1}.$$

If $\sigma \geq 1$, this follows from $\psi_{i,t_k}^K \leq C_\psi \psi_{t_k}$. If $\sigma < 1$, it follows from $\psi_{i,t_k}^K \geq \psi_{t_k}$. Since $\mu(\mathcal{L}_{t_k}) \geq \delta$,

$$\frac{r_{t_k}}{w_{t_k}} \leq C_r \psi_{t_k}^{(\sigma-1)/\sigma}$$

for some finite C_r . Therefore, for a.e. i ,

$$\frac{r_{t_k}/\psi_{i,t_k}^K}{w_{t_k}/\psi^L} = \frac{r_{t_k}}{w_{t_k}} \frac{\psi^L}{\psi_{i,t_k}^K} \leq C_r \psi_{t_k}^L \psi_{t_k}^{-1/\sigma} \rightarrow 0.$$

For all sufficiently large k , capital is strictly cheaper than labor for almost every task. Hence $\mu(\mathcal{L}_{t_k}) = 0$, contradicting $\mu(\mathcal{L}_{t_k}) \geq \delta$. Thus $\mu(\mathcal{L}_t) \rightarrow 0$. $\square$

A.2.1 Structure and Long-Run Behavior of Automation Network

We next examine how the structure of spillovers across tasks shapes the long-run composition of task production and data. In the case where $\sigma = 1/\eta$ we have a complete characterization:

Proposition A.1. If $\sigma = 1/\eta$ then for any initial stock of data D_0 and any vector f , the long-run data stock and output is as follows:

$$\lim_{t \rightarrow \infty} D_t \simeq A_t \cdot \mathcal{F}_1 \quad \text{for some } (A_t)_t \text{ s.t. } \lim_{t \rightarrow \infty} A_t = \infty.$$

where $\mathcal{F}_1$ is the right principal eigenfunction of the linear operator $\mathcal{L}$:

$$\mathcal{L} \mathcal{F}_1 = \lambda_1 \mathcal{F}_1 \quad \text{where} \quad \mathcal{L}_i D = f_i^\sigma \cdot \int W(i, k) D_k dk.$$

Proposition A.1 fully characterizes the *composition* of data and long-run output as a function of (i) degree to which task i benefits from data as captured by f_i ; and the (ii) eigenfunction centrality of task i as captured by the leading right eigenfunction $\mathcal{F}_1$.

The characterization in **Proposition A.1** allows us to show that in the long run, the composition of data must converge in proportion to a task's leading eigenfunction centrality.³⁵ Indeed, in the special case that $\mathbf{f}$ is constant, the long-run composition of data and production is exactly proportional to a continuous analog of the eigenvector centrality measure that is used widely across economics and network science (see [Golub \(2025\)](#) for a survey). When $\sigma > 1/\eta$, the amenability of task in utilizing the data from others tasks is *amplified* in equilibrium. Thus, long-run differences in the data stock should be more dispersed than is predicted by eigenfunction centrality of W . By contrast, when $\sigma < 1/\eta$, this is *dampened* and long-run differences in the data stock should be less dispersed than is predicted by the eigenfunction centrality.

Proof of Proposition A.1. Notice

$$\begin{aligned}
\lim_{t \rightarrow \infty} \frac{D_{it}}{D_{jt}} &= \lim_{t \rightarrow \infty} \frac{y_{it}}{y_{jt}} && \text{(L'Hopital's)} \\
&= \lim_{t \rightarrow \infty} \frac{\psi_{it}^K K_{it}}{\psi_{jt}^K K_{jt}} && \text{(tasks produced by capital asymptotically)} \\
&= \lim_{t \rightarrow \infty} \left(\frac{\psi_{it}^K}{\psi_{jt}^K} \right)^\sigma && (K_{it} = \frac{(\psi_{it}^K)^{\sigma-1}}{\int_0^{\gamma_t} (\psi_{jt}^K)^{\sigma-1} dj} K) \\
&= \lim_{t \rightarrow \infty} \left(\frac{f_i}{f_j} \right)^\sigma \left(\frac{\mathcal{A}_i(\mathbf{D}_t)}{\mathcal{A}_j(\mathbf{D}_t)} \right)^{\eta\sigma} \\
&= \lim_{t \rightarrow \infty} \left(\frac{f_i}{f_j} \right)^\sigma \left(\frac{\int W(i, k) D_{kt} dk}{\int W(j, k) D_{kt} dk} \right)^{\eta\sigma}
\end{aligned}$$

Hence we are looking for a data vector $\mathbf{D}$ that solves:

$$D_i = C_t \cdot f_i^\sigma \left(\cdot \int_X W(i, k) D_k dk \right)^{\eta\sigma}$$

³⁵We believe the result continues to hold in the case where $\sigma \neq 1/\eta$. Here, the operator $\mathcal{L}$ is no longer linear, and convergence requires pairing more recent results on positivity of eigenfunctions of nonlinear (convex when $\sigma > 1/\eta$; concave when $\sigma < 1/\eta$) operators with those of monotone dynamical systems.

where C is a constant independent of the task index $i \in X$. We can define the operator

$$\mathcal{L} : \mathcal{D} \rightarrow \mathcal{D} \quad \text{as} \quad \mathcal{L} \mathbf{D}_i = f_i^\sigma \left(\int_0^1 W(i, k) \mathbf{D}_k dk \right)^{\eta\sigma}$$

so we have the Fredholm integral equation of the second kind:

$$\mathbf{D} = C \cdot \mathcal{L} \mathbf{D}$$

and in the special case in which $\sigma = 1/\eta$, we have:

$$\mathbf{D}_i = C \cdot f_i^\sigma \cdot \left(\int_X W(i, k) \mathbf{D}_k dk \right)$$

and $\mathcal{L}$ is linear, compact, but not necessarily self-adjoint since spillovers are not necessarily symmetric. The spectral theorem for compact operators guarantees the existence of solutions. Substituting, we have the equation

$$\mathbf{D} = C \cdot \mathcal{L} \mathbf{D} \implies \mathcal{L} \mathbf{D} = \underbrace{\lambda}_{=: 1/C} \mathbf{D}.$$

What is C ? Equivalently, what is the corresponding eigenvalue of the graphon W ? We will show that it converges to the *principal eigenfunction* of the operator $\mathcal{L}$. We start with the following lemma that invokes a basic fact from spectral theory:

Lemma A.2. The operator $\mathcal{L}$ admits a unique principal eigenvalue λ_1 and associated eigenfunction $\mathcal{F}_1$ with strictly positive spectral gap i.e., $\text{Re}(\lambda_1 - \lambda_k) > 0$ for any $k > 1$.

This follows by observing that $\mathcal{L}$ is a positive (f_i, W are both positive), compact, and irreducible operator (since W is connected by assumption) and applying the Krein-Rutman theorem which is an analog of the Perron-Frobenius theorem for infinite-dimensional Banach spaces.

Now consider our law of motion for the data *ratio* we derived above:

$$\frac{\partial \mathbf{D}_{it} / \partial t}{\partial \mathbf{D}_{jt} / \partial t} = \frac{\mathcal{L}_i \mathbf{D}_t}{\mathcal{L}_j \mathbf{D}_t}. \quad (47)$$

This can be rearranged to separate terms by the task index:

$$\frac{1}{\mathcal{L}_i \mathbf{D}_t} \dot{\mathbf{D}}_{it} = \frac{1}{\mathcal{L}_j \mathbf{D}_t} \dot{\mathbf{D}}_{jt}$$

Since this equality must hold for all pairs $i, j \in X$, the expression must be equal to a quantity that is independent of the task index i . Call this scalar function $\alpha(t)$ (that

might depend on time):

$$\frac{1}{\mathcal{L}_i \mathbf{D}_t} \dot{\mathbf{D}}_{it} = \alpha(t)$$

This gives the following infinite-dimensional ODE for our system:

$$\dot{\mathbf{D}} = \alpha(t) \cdot \mathcal{L} \mathbf{D} \quad (48)$$

The scalar function $\alpha(t)$ in (48) affects the *speed* of evolution along the solution trajectory but does not alter the *shape* of the trajectory i.e., the composition of the data. We thus perform a time change by writing $\tau(t) = \int^t \alpha(s) ds$, which simplifies the equation to:

$$\frac{d\mathbf{D}_\tau}{d\tau} = \mathcal{L} \mathbf{D}_\tau \quad (49)$$

We solve this equation by expanding the solution in the basis of the eigenfunctions $\{\mathcal{F}_k\}_{k=1}^\infty$ of $\mathcal{L}$. Fix some large time T for which our above equation holds (taking nested limits if you prefer). Then given the data stock $\mathbf{D}_T$, since $\mathcal{L}$ is a linear operator, we have the Riesz projection onto the dominant subspace and the ‘residual’:

$$\mathbf{D}_T = \underbrace{c_1 \mathcal{F}_1}_{\substack{\text{Riesz} \\ \text{projection} \\ P\mathbf{D}_T}} + w$$

where w is the residual $(I - P)\mathbf{D}_T$ and P is the Riesz projection associated with the principal eigenvalue. Next, since $\mathcal{L}$ is time-invariant, the solution to our ODE (49) is given by the operator exponential:

$$\begin{aligned} \mathbf{D}_{\tau-\tau(T)} &= e^{\tau \mathcal{L}} \mathbf{D}_T \\ &= c_1 e^{\tau \mathcal{L}} \mathcal{F}_1 + e^{\tau \mathcal{L}} w \\ &= c_1 e^{\lambda_1 \tau} \mathcal{F}_1 + e^{\tau \mathcal{L}} w \end{aligned} \quad (\mathcal{F}_1 \text{ is an eigenfunction})$$

Next notice that we can bind the second term: define $\gamma := \sup\{\text{Re}(\lambda) : \lambda \in \sigma(\mathcal{L}), \lambda \neq \lambda_1\}$ so for any $\epsilon < \lambda_1 - \gamma$ we have:

$$\|e^{\tau \mathcal{L}} w\| \leq M e^{(\epsilon + \gamma)\tau} \|w\|.$$

Now applying [Lemma A.2](#), we have

$$\mathbf{D}_{\tau-\tau(T)} = e^{\lambda_1 \tau} \left[c_1 \mathcal{F}_1 + \underbrace{\frac{e^{\tau \mathcal{L}} w}{e^{\lambda_1 \tau}}}_{\rightarrow 0} \right]$$

so the shape of the data distribution $\mathbf{D}_t$ converges to that of the principal eigenfunction $\mathcal{F}_1$ of the operator $\mathcal{L}$:

$$\frac{D_{\tau-\tau(T)}}{\|D_{\tau-\tau(T)}\|} \rightarrow \frac{\mathcal{F}_1}{\|\mathcal{F}_1\|}$$

and we indeed have

$$\mathbf{D}_t \simeq A_t \cdot \mathcal{F}_1$$

where A_t is a time-dependent scalar tracking the aggregate data level. $\square$

A.3 Section 5: Planner's Problem

Proof of Proposition 6. The proof proceeds as follows. Throughout the proof, we will compare the planner's optimized allocation k_{it} with an auxiliary allocation that is myopically optimal given the same data stock as the planner's, $\bar{k}_{it}$. We will then compare $\bar{k}_{it}$ to the fully myopic allocation starting at $t = 0$, which we denote by k_{it}^m . We will show that these paths exhibit a *weak ordering that holds for all $t \in [0, T]$* for $\sigma \geq 1$, but not generically when $\sigma < 1$. The reason for comparing to the static myopic allocation is that the fully myopic allocation will in general feature a different data stock as the planner's for any $t > 0$.

We will first show the planner's equilibrium path coincides with the market's if and only if $\sigma = 1/\eta$. Then, we will sign the direction of bias to show that the planner leans into the market's allocation for capital. In other words, if the market allocates more capital towards unproductive tasks than productive ones, as in the complements case, the planner would do so even more aggressively.

Generic inefficiency. The planner's current value Hamiltonian is

$$\mathcal{H}(\mathbf{k}_t, \mathbf{D}_t, \boldsymbol{\lambda}_t, \mu_t) = Y_t + \sum_{i=1}^2 \lambda_{it} y_{it} + \mu_t \left(K - \sum_{i \in \{h, l\}} \frac{1}{2} k_{it} \right)$$

We will take i to mean the h task and j to mean the l task. The FOCs are

$$\begin{aligned} \frac{1}{2} Y_t^{\frac{1}{\sigma}} y_{it}^{-\frac{1}{\sigma}} \psi^K(D_{it}) + \lambda_{it} \psi^K(D_{it}) - \frac{1}{2} \mu_t &= 0 \\ \frac{1}{2} Y_t^{\frac{1}{\sigma}} y_{it}^{-\frac{1}{\sigma}} (\psi^K(D_{it}))' k_{it} + \lambda_{it} (\psi^K(D_{it}))' k_{it} &= \rho \lambda_{it} - \dot{\lambda}_{it} \end{aligned}$$

The planner's relative task production is thus

$$\frac{y_{it}}{y_{jt}} = \left(\frac{\frac{1}{2} \frac{\mu_t}{\psi^K(D_{jt})} - \lambda_{jt}}{\frac{1}{2} \frac{\mu_t}{\psi^K(D_{it})} - \lambda_{it}} \right)^\sigma$$

We know in the myopic case, given the same data stock, the ratio would be

$$\frac{\bar{y}_{it}}{\bar{y}_{jt}} = \left(\frac{\psi^K(D_{it})}{\psi^K(D_{jt})} \right)^\sigma$$

So the planner allocates more capital to task i compared to the myopic allocation with the same data stock if and only if

$$\lambda_{jt} \psi^K(D_{jt}) \leq \lambda_{it} \psi^K(D_{it})$$

We know the relationship holds with equality when $t = T$ since $\lambda_{iT} = \lambda_{jT} = 0$. We will evolve the system backwards in time. Differentiating the RHS, we have

$$\begin{aligned} (\psi_{it}^K)' y_{it} \lambda_{it} + \psi_{it}^K \dot{\lambda}_{it} &= (\psi_{it}^K)' y_{it} \lambda_{it} + \psi_{it}^K \left(\rho \lambda_{it} - \frac{1}{2} Y_t^{\frac{1}{\sigma}} y_{it}^{-\frac{1}{\sigma}} (\psi_{it}^K)' k_{it} - \lambda_{it} (\psi_{it}^K)' k_{it} \right) \\ &= (\psi_{it}^K)' y_{it} \lambda_{it} + \psi_{it}^K \left(\rho \lambda_{it} - \left(\frac{1}{2} Y_t^{\frac{1}{\sigma}} y_{it}^{-\frac{1}{\sigma}} + \lambda_{it} \right) (\psi_{it}^K)' k_{it} \right) \\ &= (\psi_{it}^K)' y_{it} \lambda_{it} + \psi_{it}^K \rho \lambda_{it} - \frac{1}{2} \mu_t (\psi_{it}^K)' k_{it} \end{aligned}$$

The LHS is symmetric. The two sides are equal at $t = T$ if and only if

$$\begin{aligned} (\psi_{jt}^K)' y_{jt} \lambda_{jt} + \psi_{jt}^K \rho \lambda_{jt} - \frac{1}{2} \mu_t (\psi_{jt}^K)' k_{jt} &= (\psi_{it}^K)' y_{it} \lambda_{it} + \psi_{it}^K \rho \lambda_{it} - \frac{1}{2} \mu_t (\psi_{it}^K)' k_{it} \\ (\psi_{jt}^K)' k_{jt} \left(\psi_{jt}^K \lambda_{jt} - \frac{1}{2} \mu_t \right) &= (\psi_{it}^K)' k_{it} \left(\psi_{it}^K \lambda_{it} - \frac{1}{2} \mu_t \right) \quad (\Longleftrightarrow) \\ \frac{(\psi_{jt}^K)'}{(\psi_{it}^K)'} &= \frac{k_{it}}{k_{jt}} \quad (\Longleftrightarrow) \\ \left(\frac{\psi_{it}^K}{\psi_{jt}^K} \right)^{1/\eta} &= \left(\frac{\psi_{it}^K}{\psi_{jt}^K} \right)^\sigma \quad (\Longleftrightarrow) \end{aligned}$$

When $\sigma = 1/\eta$, the guess holds, and the planner indeed allocates capital according to her myopic optimal allocation given her current data stock. Since this holds at all times, we conclude $k_{it} = \bar{k}_{it} = k_{it}^m$ for $i = l, h$ and t . When $\sigma \neq 1/\eta$, however, this can only hold if $D_{it} = D_{jt}$, but if we evolve the system back to $t = 0$, this contradicts the initial condition that one task begins with strictly higher data. As such, we conclude the paths will be *inefficient* when $\sigma \neq 1/\eta$. Moreover, from the derivation above, we note that the RHS is greater than the LHS for some $t = T - \epsilon$ with $\epsilon > 0$ sufficiently

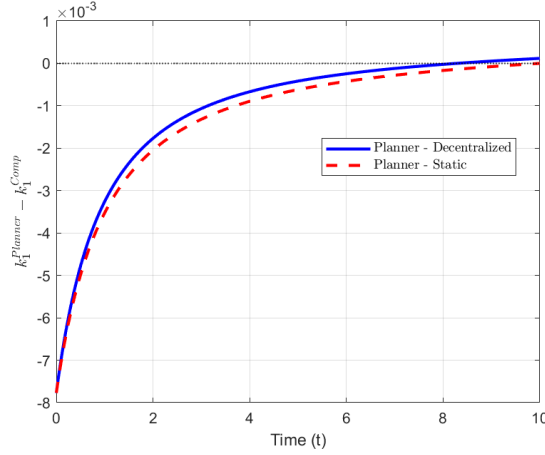


Figure 15: Comparison of capital allocated to the data-rich sector.

small if and only if

$$\left(\frac{\psi_{it}^K}{\psi_{jt}^K} \right)^{1/\eta} \leq \left(\frac{\psi_{it}^K}{\psi_{jt}^K} \right)^\sigma$$

where we use the fact that $\psi_{jt}^K \lambda_{jt} - \frac{1}{2} \mu_t < 0$ and that we are running the system *backward* in time. That is, provided $\psi_{it}^K > \psi_{jt}^K$, the planner produces more of i relative to the myopic optimal allocation given her data stock if and only if $\sigma > 1/\eta$. In the next part of this argument, we extend this local result near T to a global result regarding the entire path, and we will compare with the fully myopic path instead of the myopic path given the planner's data stock.

Global comparison. Towards making a global comparison, we will first show [Lemma A.3](#), which states that either $k_{ht} \geq \bar{k}_{ht} \forall t$ or $k_{ht} \leq \bar{k}_{ht} \forall t \in [0, T]$. Put differently, the planner's allocation and her myopically optimal allocation cannot cross. Using this result, we can now conclude the proof.

1. $\sigma < 1$. In this case, we must have $k_{ht} \leq \bar{k}_{ht} \forall t$ and strictly so for at least some interval by combining our earlier argument with [Lemma A.3](#).³⁶ Additionally, note that $\bar{k}_{h0} = k_{h0}^m$ since data stocks at $t = 0$ are equivalent, so $k_{h0}^m \leq \bar{k}_{h0}$. However, we do not necessarily have $k_{ht} \leq k_{ht}^m$ for all times as the productivity of the initially data poor task is growing faster under the planner's allocation. Given the planner's initial tilt towards the low data task, we may reach a stage where the planner balances her allocation of capital across tasks sooner compared to the decentralized equilibrium, so we may very well have $k_{ht} > k_{ht}^m$ at some point.

³⁶We cannot have the relation always hold with equality. Otherwise, we would contradict that one task has strictly higher data to begin with as argued previously.

Figure 15 illustrates this reversal numerically: whereas the planner allocates less capital to the data-poor sector compared to the decentralized equilibrium to begin with, the allocation reverses around $t \approx 8$.

2. $1 \leq \sigma < 1/\eta$. As before, we have $k_{ht} \leq \bar{k}_{ht} \forall t$. We now compare $\bar{k}_{ht}$ with k_{ht}^m . The two are initially equal. Then the planner allocates weakly less capital compared to the decentralized equilibrium, thereby accumulating less data, and so on and so forth. It is thus straightforward to verify that a lower ratio of data (of h to l) implies less capital allocated to h when $\sigma \geq 1$ at all times. Therefore we have the following ordering: $k_{ht} \leq \bar{k}_{ht} \leq k_{ht}^m$.
3. $\sigma = 1/\eta$. This is covered by the efficiency result from before.
4. $\sigma > 1/\eta$. In this case, we must have $k_{ht} \geq \bar{k}_{ht} \geq k_{ht}^m \forall t$ and the first inequality holding strictly at least for some interval. The argument is analogous to the second case.

□

Lemma A.3 (No crossing). Either $k_{ht} \geq \bar{k}_{ht} \forall t$ or $k_{ht} \leq \bar{k}_{ht} \forall t \in [0, T]$.

Proof of Lemma A.3. Towards a contradiction, suppose there exists some $\tau \in (0, T)$ where k_{ht} crosses from below to above of $\bar{k}_{ht}$. The other direction proceeds analogously. Intuitively, the planner first over invests in the low data task and then undoes her investment, which cannot be optimal. To establish this rigorously, we will construct a perturbation by shifting capital between l and h over a small interval around τ . In doing so, we do not affect output prior to the perturbation and strictly improve the objective over this short interval. Our construction will also have the property that the perturbed data vector after the interval coincides with the data vector from the planner's original equilibrium, so the outputs are equal after the interval as well. Clearly, such a perturbation strictly improves the objective, contradicting optimality.

It remains to show that such a construction exists. Let $D_i^p(s)$ denote the planner's data of i at time s ; let $\tilde{D}_i(s)$ be the data of i at time s following the perturbed capital path. We make the time dependence more explicit for clarity. Fix some $\epsilon > 0$. Formally, our feasibility requirements are

$$\begin{aligned}
D_i^p(\tau - \epsilon) &= \tilde{D}_i(\tau - \epsilon) \\
D_i^p(\tau + \epsilon) &= \tilde{D}_i(\tau + \epsilon) \\
\dot{\tilde{D}}_1(s) &= f(\tilde{D}_1(s))^\eta \left(k_1(s) + \Delta_1^k(s) \right) \quad s \in [\tau - \epsilon, \tau + \epsilon] \\
\dot{\tilde{D}}_2(s) &= f(\tilde{D}_2(s))^\eta \left(k_2(s) - \Delta_1^k(s) \right) \quad s \in [\tau - \epsilon, \tau + \epsilon]
\end{aligned}$$

Here $f \equiv f_1 = f_2$. From ODE theory, we recognize that $\tilde{D}_i(s)$ solve a Bernoulli differential equation which admits closed form solutions. Our requirement of reaching the same data vector at $\tau + \epsilon$ boils down to

$$D_1^p(\tau + \epsilon) = \left((\tilde{D}_1(\tau - \epsilon))^{1-\eta} + (1-\eta) f \int_{\tau-\epsilon}^{\tau+\epsilon} \left(k_1(s) + \Delta_1^k(s) \right) ds \right)^{\frac{1}{1-\eta}}$$

$$D_2^p(\tau + \epsilon) = \left((\tilde{D}_2(\tau - \epsilon))^{1-\eta} + (1-\eta) f \int_{\tau-\epsilon}^{\tau+\epsilon} \left(k_2(s) - \Delta_1^k(s) \right) ds \right)^{\frac{1}{1-\eta}}$$

Observe that for any Δ_1^k that is anti-symmetric around τ , the above relations must hold. As such, simply choose a $\Delta_1^K : [\tau - \epsilon, \tau + \epsilon] \rightarrow \mathbb{R}_{>0}$ satisfying

$$\begin{cases} k_1(\tau - s) < k_1(\tau - s) + \Delta_1^K(\tau - s) < \bar{k}_1(\tau - s) & \forall s \in [0, \epsilon] \\ \bar{k}_1(\tau + s) < k_1(\tau + s) + \Delta_1^K(\tau + s) < k_1(\tau + s) & \forall s \in [0, \epsilon] \end{cases}$$

For example, we can choose a linear segment anchored at $k_1(\tau) = \bar{k}_1(\tau)$ that is above the planner's allocation but sufficiently close to it between $[\tau - \epsilon, \tau]$. Call the difference $\Delta_1^K(\tau - s)$. We would simply take the negative of the difference for $[\tau, \tau + \epsilon]$, which would lie in the interval $(k_1^M(\tau + s), k_1^M(\tau + s))$ provided the initial segment is close enough to the planner's allocation. $\square$

A.4 Section 6: Capital Accumulation

In this section, we first derive the singularity result under a Solow capital accumulation rule; we then bound the savings appropriately in a neoclassical growth model and apply our Solow result to conclude.

With Solow capital accumulation, again, we have

$$1 - \gamma_t \equiv \chi_t = \frac{L}{K} D_t^{-\eta}$$

$$\dot{D}_t = \frac{L}{\chi_t}$$

The evolution of capital is given by

$$\begin{aligned} \dot{K}_t &= -\delta K_t + s_t Y_t \\ &= -\delta K_t + s_t \left(\gamma_t^{\frac{1}{\sigma}} (D_t^\eta K_t)^{\frac{\sigma-1}{\sigma}} + (1 - \gamma_t)^{\frac{1}{\sigma}} L^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \end{aligned}$$

Lemma A.4. When sectors are symmetric, and the savings rate satisfies

$$\inf_{t \geq T} s_t \geq s > 0$$

for some T finite, then there is a finite time explosion.

Proof of Lemma A.4. The proof proceeds as follows. First, we will run the system long enough so that $Y_t \geq D_t^\eta K_t \geq \delta K_t$. This ensures capital will be increasing at all times moving forward. Next, we will lowerbound capital accumulation, output, and automation by studying an economy where we ignore labor altogether. We finally establish a finite time explosion in this economy and thereby concluding the argument.

1. Finding a time such that capital always increases thereafter We will first find a t_0 such that capital is guaranteed to increase thereafter. We do this in two steps. First, notice that

$$\begin{aligned} Y_t &= D_t^\eta \frac{K_t}{\gamma_t} \\ &\geq D_t^\eta K_t \end{aligned}$$

The first equality follows from γ_t solving the following relationship

$$Y_t = Y_{i,t} = D_t^\eta \frac{K_t}{\gamma_t} = \frac{\psi^L L}{1 - \gamma_t} \quad \forall i \in X$$

Next, notice that

$$\dot{D}_t = Y_t = \frac{\psi^L L}{1 - \gamma_t} = \psi^L L > 0$$

This implies D_t tends to infinity, so we can find a t_0 such that $D_{t_0}^\eta K_{t_0} > \delta K_{t_0} \implies \dot{K}_{t_0} > 0$. Then, by virtue of D_t being strictly increasing in time, $\dot{K}_t > 0$ for all $t \geq t_0$.

2. A lower-bounded economy We next consider $t \geq \max\{T, t_0\} \equiv \underline{t}$ by studying a simplified counterfactual economy. The actual economy evolves according to

$$\begin{cases} \dot{K}_t &= -\delta K_t + s_t D_t^\eta \frac{K_t}{\gamma_t} \\ \dot{D}_t &= D_t^\eta \frac{K_t}{\gamma_t} \end{cases}$$

If we produce with capital only and a lower bounded savings rate s , the dynamics are

$$\begin{cases} \dot{\tilde{K}}_t &= -\delta \tilde{K}_t + s \tilde{D}_t^\eta \tilde{K}_t \\ \dot{\tilde{D}}_t &= \tilde{D}_t^\eta \tilde{K}_t \end{cases}$$

By a version of Kamke's Comparison Theorem (Smith, 1995), we can show $K_t \geq \tilde{K}_t$ and $D_t \geq \tilde{D}_t$ for all $t \geq t_0$ if we specialize the two economies to begin with the same initial conditions at $\underline{t}$ as the system is cooperative.

3. Dynamics of the lower-bounded economy Now we turn to the dynamics of the lower-bounded economy at $t \geq \underline{t}$. The capital evolution is given by

$$\dot{\tilde{K}}_t = -\delta \tilde{K}_t + s \tilde{D}_t^\eta \tilde{K}_t$$

Since D_t is going to infinity in the limit, we can choose t_1 so that for some $0 < \epsilon < s$

$$\dot{\tilde{K}}_{t_1} = -\delta \tilde{K}_{t_1} + s \tilde{D}_{t_1}^\eta \tilde{K}_{t_1} \geq (s - \epsilon) \tilde{D}_{t_1}^\eta \tilde{K}_{t_1}$$

Again, by virtue of D_t being increasing in t , the above relation holds for all $t \geq t_1$. From the data LOM in this economy, we have

$$\dot{\tilde{K}}_t \geq (s - \epsilon) \dot{\tilde{D}}_t$$

Integrating from t_1 to t , the relationship is

$$\tilde{K}_t \geq (s - \epsilon) \tilde{D}_t - \tilde{D}_{t_1} + \tilde{K}_{t_1}$$

We can then find some t_2 large enough so that for some $\omega > 0$

$$\tilde{K}_{t_2} \geq (s - \epsilon) \tilde{D}_{t_2} - \tilde{D}_{t_1} + \tilde{K}_{t_1} \geq \omega \tilde{D}_{t_2}$$

This relationship similarly holds for all $t \geq t_2$ as $\tilde{D}_t$ is monotonically increasing. Substituting this relation back into the dynamics for data accumulation, we have for all $t \geq t_2$

$$\dot{\tilde{D}}_t \geq \omega \tilde{D}_t^{1+\eta}$$

This super-linear growth implies a finite time explosion. In particular, we know that

$$\tilde{D}_t \geq \left(\tilde{D}_{t_2}^{-\eta} - \omega \eta (t - t_2) \right)^{-\frac{1}{\eta}}$$

which means full automation as well as the finite time explosion in data and output takes place weakly before $t_2 + 1 / (\omega \eta \tilde{D}_{t_2}^\eta)$ in the original economy. $\square$

Proof of Proposition 7. We now move to the NGM set-up. The representative household solves

$$\max_{\{c_t\}} \int_0^\infty e^{-\rho t} \frac{c_t^{1-\theta}}{1-\theta} dt \quad \text{s.t.} \quad \dot{a}_t = r_t a_t + w_t - c_t$$

yielding the Euler equation

$$\frac{\dot{c}_t}{c_t} = \frac{1}{\theta} (r_t - \delta - \rho)$$

and the aggregate resource constraint $\dot{K}_t = s_t Y_t - \delta K_t$, where $s_t = (Y_t - c_t) / Y_t$ is the equilibrium savings rate. To apply the previous result, we want to show that the savings rate $s_t \rightarrow s > 0$, so the savings rate ought to be lower bounded for some $t = T$ sufficiently large.

We proceed by guess-and-verify and let t^* denote the time of singularity.³⁷ Assume $s_t \rightarrow s > 0$, then for t sufficiently large

$$\dot{K}_t = s_t Y_t - \delta K_t \geq \omega_t Y_t$$

Where ω_t is chosen as in the earlier proof, and $\lim_{t \rightarrow t^*} \omega_t = s$. We also have $\dot{D}_t = Y_t$. As such, $\lim_{t \rightarrow t^*} K_t / D_t = s_t$ as we approach the singularity, which implies the growth rates of these variables, $g_{K,t}$ and $g_{D,t}$, must also converge. From the resource constraint,

$$g_{Y,t} = \eta g_{D,t} + g_{K,t} \approx (1 + \eta) s D_t^\eta$$

From the Euler equation,

$$g_{c,t} \approx \frac{1}{\theta} D_t^\eta$$

By virtue of the guess, consumption and output growth rate must equalize, so

$$s = \frac{1}{(1 + \eta)\theta}$$

This completes the guess-and-verify step. Applying [Lemma A.4](#) establishes that output and consumption will become unbounded in finite time, and the entire economy will be automated in finite time as well. Finally, when $\theta \leq 1$, the candidate path is feasible and delivers infinite expected utility; when $\theta > 1$, the transversality condition is satisfied. This shows the candidate path is both feasible and optimal, concluding our proof. $\square$

B Extensions

A multisector economy. In our economy tasks are aggregated directly into a final good. We might, instead, consider an economy comprised of a set of intermediate goods (M) and a spectrum of tasks (X), in which a final good is produced with intermediate goods, and each intermediate good is produced with a bundle of tasks.

³⁷If helpful, this can be recast as a proof by contradiction, so $t^* = \infty$.

Different intermediate goods value different tasks—the degree to which intermediate good $m \in M$ values task $i \in X$ is $G(m, i) \geq 0$. Intermediate goods overlap in the set of tasks that they value, so data of task i is generated by and used in production of all the intermediate goods that use task i .³⁸

Appendix B.1 outlines such an economy and sketches how relative elasticities for the final good vis-a-vis intermediate goods, as well as how the pattern of intermediate good-task dependance and how they overlap, shapes differential data accumulation.

Non-substitutable spillovers. Our model of data spillovers imposed the implicit assumption that although the importance of task j on i ’s productivity is weighted by $W(i, j)$, they still enter into i ’s production function in a *substitutable* fashion. While this is standard in economic analyses of networks, we might imagine richer data aggregation in which we require both data from task i *and* task j to augment capital productivity for i :

$$\left(\int \left(W(i, j) D_{jt} \right)^{\frac{\phi-1}{\phi}} dj \right)^{\frac{\phi}{\phi-1}} \quad \text{where } \phi \in [0, +\infty).$$

Our model of cross-task spillovers corresponds to $\phi \rightarrow +\infty$; in the other extreme $\phi \rightarrow 0$ data-driven automation is then held up by the data-poorest task.³⁹ Of course, this ‘CES-form’ imposes a kind of symmetry in how task data aggregates; we think a richer model capturing higher-order interactions among task-specific data is an exciting line of future work.

Expanding the set of tasks. Although our main analysis holds the task space fixed, our model offers a novel perspective on why human labor might have a persistent advantage over capital at new tasks. Let $X_t = [0, x_t]$ denote the set of tasks used in the final good production where the *task frontier* x_t is continuous and increasing in t . For a new task x_t that is introduced at time- t , it is, by definition, novel and was not previously produced so at the time t at which it was introduced we have no task-specific data. How might we expect automation and data accumulation in this economy to unfold?

The main trade-off stems from the degree of spillover among the new task and other exiting tasks. If the spillover is weak, labor will have a comparative advantage when the new task is introduced, although the advantage might not persist. In this econ-

³⁸The idea that different sectors might be connected by sharing ‘overlapping tasks’ that precipitates data spillovers is related in spirit to [Chen, Elliott, and Koh \(2023\)](#) in which different ‘markets’ share ‘overlapping capabilities’ that precipitates merger waves.

³⁹For instance if $W(i, \cdot)$ is full-support on X this slows down the transition path of automation substantially since the rate at which effective data grows is held-up by, e.g., the most data-poor sector. This has sharp implications for the planner’s problem of how to allocate production across sectors.

omy, *human labor facilitates automation* by producing frontier tasks—and attendant training data—precisely when data is scarce.⁴⁰ Alternatively, if there is sufficient overlap with existing tasks (as given by our graphon W that captures cross-task spillovers), task x_t can be automated by training AI systems on nearby tasks. In this world, strong data spillovers allow automation to *bootstrap itself* by continually using the endogenously generated data on tasks at the frontier to automate new tasks as soon as they emerge.

Which of these economies emerge will ultimately depend on the *speed* at which new tasks emerge relative to the *strength* of spillovers from existing tasks to new tasks. When new tasks emerge quickly, tasks near the frontier, have, themselves, not been around for long—and so they are “data poor”. Then, in the absence of strong spillovers from the core tasks, labor will play an important dual role in both producing new tasks as well as generating data on them. Alternatively, When new tasks emerge slowly, tasks near the frontier have been around for long enough that they are “data rich”, and even local spillovers may well suffice for the new task to be rapidly automated.

B.1 A Multi-sector Economy

In the main text we focused on a single-sector economy that aggregated tasks into a final good. Here we extend our main model to one with multiple sectors that is aggregated via another CES into a final good. All sectors draw from the same continuum of tasks, but different sectors may use the same task at different intensities, or combine tasks differently via CES aggregators with different task substitutabilities.

Tasks and the final good. We now index intermediate products by $m \in M := [0, 1]$. A final good is now produced competitively

$$Q = \left(\int_M q_m^{\frac{v-1}{v}} \right)^{\frac{v}{v-1}}$$

where q_m is the quantity of sector m produced competitively.

Intermediate products. Intermediate products are produced by tasks:

$$q_m = \left(\int_X (G(m, i) \cdot y_i^m)^{\frac{\sigma_m-1}{\sigma_m}} di \right)^{\frac{\sigma_m}{\sigma_m-1}}.$$

⁴⁰For instance, in [Korinek and Suh \(2024\)](#) it is assumed that humans can simply perform high complexity tasks.

where, as before, $X = [0, 1]$ is our task space and y_i^m is the quantity of task $i \in X$ used to produce sector $m \in M$. The function

$$G(m, \cdot) : X \rightarrow \mathbb{R}_+$$

governs the degree to which task $i \in X$ is used in the production of sector $m \in M$.

Remark 2. The analysis in the main text is thus a special case where the economy was comprised of a single sector m , and $G(m, i) = 1$ for each $i \in X$.

Tasks. Task i can be performed with capital or labor:

$$y_i^m = \psi^L \cdot L_i^m + \psi^K(D_i) \cdot K_i^m$$

where L_i^m, K_i^m are the quantities of labor and capital employed in sector m to produce task i . We assume, as in the main text, that

$$\psi^K(\mathcal{A}_i(\mathbf{D})) = (\mathcal{A}_i(\mathbf{D}))^\eta \quad \text{for } \eta \in (0, 1).$$

and to simplify our exposition we assume $\mathcal{A}_i(\mathbf{D}) = D_i$ i.e., no cross-task spillovers.

Law of motion aggregated across sectors. We have the law of motion

$$\dot{D}_i = \int_M y_i^m dm \quad \text{for each task } i \in X$$

i.e., the evolution of the data for task i depends on the production of *all* sectors.

Task cost. The resulting unit cost to produce using task i given data stock $\mathbf{D}$ is:

$$c_i(D_i) = \min \left(\frac{w}{\psi^L}, \frac{r}{\psi^K(D_i)} \right) = \min \left(\frac{w}{\psi^L}, \frac{r}{D_i^\eta} \right) \quad (50)$$

Intermediate product cost. The price index for intermediate product m is its unit cost function:

$$p_m(\mathbf{D}) = \left(\int_{j \in X} \left(\frac{c_j(D_j)}{G(m, j)} \right)^{1-\sigma_m} dj \right)^{\frac{1}{1-\sigma_m}}. \quad (51)$$

where $\sigma_m > 0$ is the CES parameter for sector $m \in M$.

Final good production and demand for intermediate products. The demand for sector m 's output is given by the usual FOC condition for the final good producer:

$$q_m(D) = Q \left(\frac{p_m(\mathbf{D})}{P} \right)^{-\nu} \quad (52)$$

where P is the aggregate price index of the final good Q . We will set the final good as the numeraire in this economy so $P = 1$.

Task demand. Task demand for intermediate product m can be written:

$$y_i^m(\mathbf{D}) = q_m \cdot (G(m, i))^{\sigma_m - 1} \cdot \left(\frac{c_i}{p_m} \right)^{-\sigma_m}$$

and so aggregate task demand can be written:

$$\begin{aligned} y_i(\mathbf{D}) &= \int_{m \in M} y_i^m(\mathbf{D}) dm \\ &= \int_{m \in M} \left[q_m \cdot (G(m, i))^{\sigma_m - 1} \cdot \left(\frac{c_i}{p_m} \right)^{-\sigma_m} \right] dm \\ &= \int_{m \in M} \left[Q(p_m(\mathbf{D}))^{-\nu} (G(m, i))^{\sigma_m - 1} \cdot \left(\frac{c_i}{p_m} \right)^{-\sigma_m} \right] dm \\ &= Q \int_{m \in M} \left[(p_m(\mathbf{D}))^{\sigma_m - \nu} (G(m, i))^{\sigma_m - 1} \cdot (c_i)^{-\sigma_m} \right] dm \\ &= Q \int_{m \in M} \underbrace{\left(\int_{j \in X} \left(\frac{c_j(\mathbf{D}_j)}{G(m, j)} \right)^{1 - \sigma_m} dj \right)^{\frac{\sigma_m - \nu}{1 - \sigma_m}}}_{= p_m^{\sigma_m - \nu}} (G(m, i))^{\sigma_m - 1} \cdot (c_i)^{-\sigma_m} dm \end{aligned}$$

We make three key observations about how the quantity of task i is determined. This is the key variable that will, in turn, drive the corresponding task-specific data accumulation of $\mathbf{D}_i$.

First, and most straightforwardly, y_i is decreasing in c_i , albeit weighted by sectoral elasticities $(\sigma_m)_m$ inside the integral. If $\sigma_m = \sigma$ then we have a constant elasticity of the task exactly given by σ .

Second, how sectoral prices $(p_m)_m$ propagate to task quantities hinge on the sizes of $\sigma_m \geq \nu$. If $\sigma_m > \nu$ i.e., sectoral production via tasks is more substitutable than the final good production—this perhaps the economically relevant case—then an (exogenous) increase in p_m that holds c_i fixed e.g., perturbing the cost of other tasks leads to *more* of task i being produced. This is driven by two effects. First, the final goods producer demands less of sector m since

$$q_m = Q p_m^{-\nu}.$$

But as the sector price increases, there is also a task reallocation effect within sector m : producers reallocate production toward task i as long as the sector's production function is sufficiently substitutable, and this is exactly the condition $\sigma_m > \nu$.

Third, $G(m, i)$ measures the ‘importance’ of task i for sector m . If production in sector m are gross substitutes i.e., $\sigma_m \geq 1$ then when demand for sector m increases, more important tasks—those with high $G(m, i)$ —are *overproduced*, and the extent of this overproduction is given by $\sigma_m \geq 1$. Conversely, if $\sigma < 1$ then the opposite obtains: when demand for sector m increases, the less important tasks are overproduced.

C Simulation Method

We solve a discretized version of our model with a continuum of tasks. [Appendix C.1](#) introduces the discretized model and [Appendix C.2](#) outlines the numerical procedure. Time subscripts are omitted for brevity. Replication codes are available upon request.⁴¹

C.1 Discretized Model

We approximate the continuum of tasks by N large. Define $\epsilon := 1/N$, output is given by⁴²

$$Y := \left(\sum_{i=1}^N (y_i)^{\frac{\sigma-1}{\sigma}} \epsilon \right)^{\frac{\sigma}{\sigma-1}}$$

Costs are now $\sum_{i=1}^N p_i y_i \epsilon$, so the FOC for the final good producer is

$$Y^{\frac{1}{\sigma}} y_i^{-\frac{1}{\sigma}} = p_i = \min \{w/\psi_L, r/\psi_i^K\}$$

Intermediate producers still solve

$$\max_{L_i, K_i} p_i (\psi_i^K K_i + \psi^L L_i) - r_i K_i - w L_i$$

To determine the equilibrium allocation, we define an iterative procedure. Let $\tilde{\gamma}$ be the largest sector index where producing with capital is *strictly* cheaper than labor:

$$\tilde{\gamma} := \max \left\{ i \in \{0, \dots, N\} : \frac{r}{\psi_i^K} < \frac{w}{\psi^L} \right\}$$

Consequently, tasks $\leq \tilde{\gamma}$ are produced entirely by capital. We then examine the *next* sector, $\tilde{\gamma} + 1$. If production costs equalize at this sector, it becomes the “split” sector where the intermediate good producer is indifferent between producing using labor

⁴¹Please email bryantxia2435@gmail.com

⁴²We set $N = 1000$ in our codes.

versus capital. Let $\alpha \in [0, 1]$ be the fraction of task $\tilde{\gamma} + 1$ produced using capital; below we detail the procedure for solving this ratio.

Factor prices and allocations given $\tilde{\gamma}$ and α : For labor, tasks $\tilde{\gamma} + 2, \dots, N$ and the labor-portion of task $\tilde{\gamma} + 1$ must clear:

$$\epsilon \left(\sum_{i=\tilde{\gamma}+2}^N L_i + (1 - \alpha) L_{\tilde{\gamma}+1} \right) = L$$

From the final good producer's FOC and the fact that labor productivity is equalized across all tasks, L_i must be constant:

$$L_i = \frac{L}{\epsilon (N - (\tilde{\gamma} + 1) + 1 - \alpha)} = \frac{L}{\epsilon (N - \tilde{\gamma} - \alpha)}$$

For capital, tasks $1, \dots, \tilde{\gamma}$ and the capital-portion of task $\tilde{\gamma} + 1$ must clear:

$$\epsilon \left(\sum_{i=1}^{\tilde{\gamma}} K_i + \alpha K_{\tilde{\gamma}+1} \right) = K$$

Using the optimal allocation rule $K_i \propto (\psi_i^K)^{\sigma-1}$:

$$K_i = \frac{(\psi_i^K)^{\sigma-1}}{\sum_{m=1}^{\tilde{\gamma}} (\psi_m^K)^{\sigma-1} + \alpha (\psi_{\tilde{\gamma}+1}^K)^{\sigma-1}} \frac{K}{\epsilon}$$

Plugging these allocations into the final good producer's FOC, we have that factor prices conditional on $\tilde{\gamma}$ and α are given by:

$$r(\tilde{\gamma}, \alpha) = Y^{\frac{1}{\sigma}} \left(\frac{\epsilon \left(\sum_{m=1}^{\tilde{\gamma}} (\psi_m^K)^{\sigma-1} + \alpha (\psi_{\tilde{\gamma}+1}^K)^{\sigma-1} \right)}{K} \right)^{\frac{1}{\sigma}}$$

$$w(\tilde{\gamma}, \alpha) = Y^{\frac{1}{\sigma}} \left(\frac{\epsilon (N - \tilde{\gamma} - \alpha) (\psi^L)^{\sigma-1}}{L} \right)^{\frac{1}{\sigma}}$$

Solving Procedure: First, consider the set

$$\mathcal{C} := \left\{ j : \frac{\psi_j^K}{\psi_L} > \frac{r(j, 0)}{w(j, 0)} \right\}$$

If this set is empty, set $\tilde{\gamma} = 0$. Else, set $\tilde{\gamma} := \max \mathcal{C}$. Next, solve numerically for $\alpha \in [0, 1]$

such that

$$\frac{\psi_{\tilde{\gamma}+1}^K}{\psi^L} = \frac{r(\tilde{\gamma}, \alpha)}{w(\tilde{\gamma}, \alpha)}$$

C.2 Evolution Equations

At every t , recall the state variable is the vector of data, $\mathbf{D}$. We first construct ψ from data,

$$\psi_i^K(D) = f\left(\frac{i}{N}\right) \left(\sum_{m=1}^N g(|m-i|\epsilon) D_m \right)^\eta$$

Note that setting $g(0) = 1$ and 0 everywhere else corresponds to the no spillover case. Next, we determine $\tilde{\gamma}$ and α using the procedure above. The evolution of data depends on task output y_i .

- For $i \leq \tilde{\gamma}$ (Strictly Capital): $y_i = \psi_i^K K_i$.
- For $i > \tilde{\gamma} + 1$ (Strictly Labor): $y_i = \psi^L L_i$.
- For the split sector $i = \tilde{\gamma} + 1$:
 - If $\alpha = 0$: $y_i = \psi^L L_i$ (Transitions to labor).
 - If $\alpha > 0$: Costs equalize, so $p_i = w/\psi^L$. Demand implies $y_i = \psi^L L_{labor}$, where L_{labor} is the intensity L_i calculated for fully labor sectors.⁴³

Combining these, the data evolution is:

$$\frac{dD_i}{dt} = y_i - \rho D_i$$

Finally, for Solow-style capital accumulation devoid of household optimization, we can augment the simulation by including the usual capital accumulation rule and keeping track of $K(t)$

$$\dot{K} = sY - \delta K$$

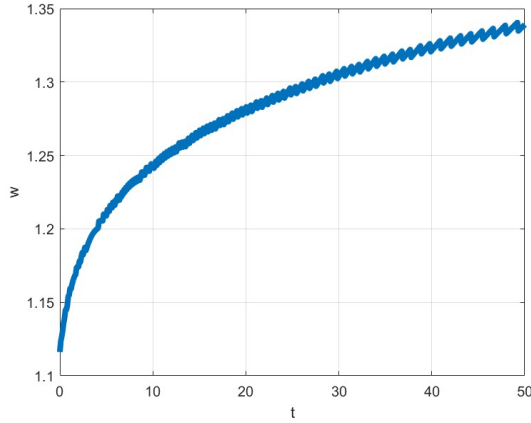
We simulate this system in Matlab.

D Additional Numerical Results

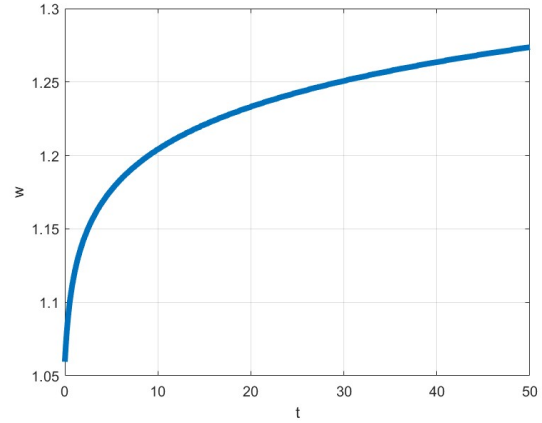
We collect additional numerical results in this appendix.

⁴³Since labor productivity for all sectors are the same, the combined capital and labor output for this sector must be equal to the output in labor only sectors.

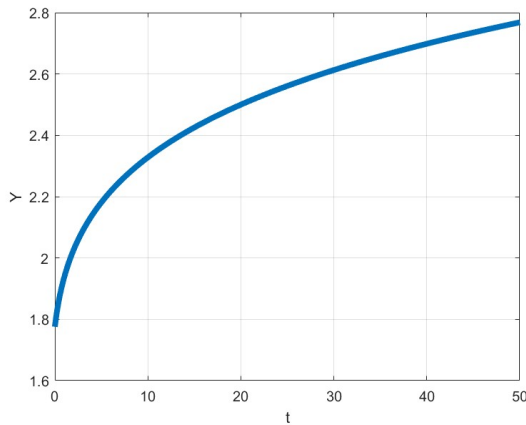
D.1 Time Paths Wages and Outputs without Capital Accumulation



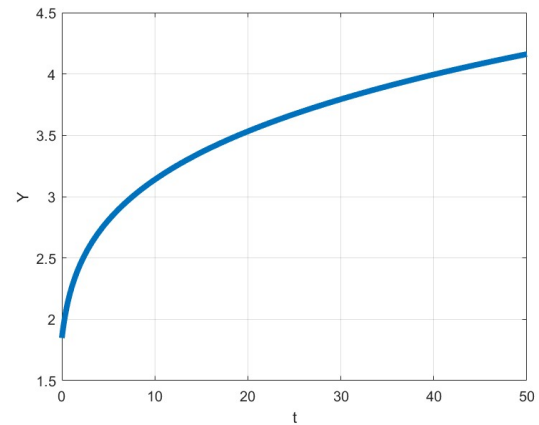
(a) Evolution of w for $\sigma = 0.5$



(b) Evolution of w for $\sigma = 5.5$



(c) Evolution of Y for $\sigma = 0.5$



(d) Evolution of Y for $\sigma = 5.5$

Figure 16: The time-path of wages and output. In the complements case, wages rise due to increased capital productivity, but this effect will eventually be fully offset by the displacement effect as labor is restricted endogenously to a shrinking set of tasks.

D.2 Peripheral Automation Across Substitutability and Bridge Strength

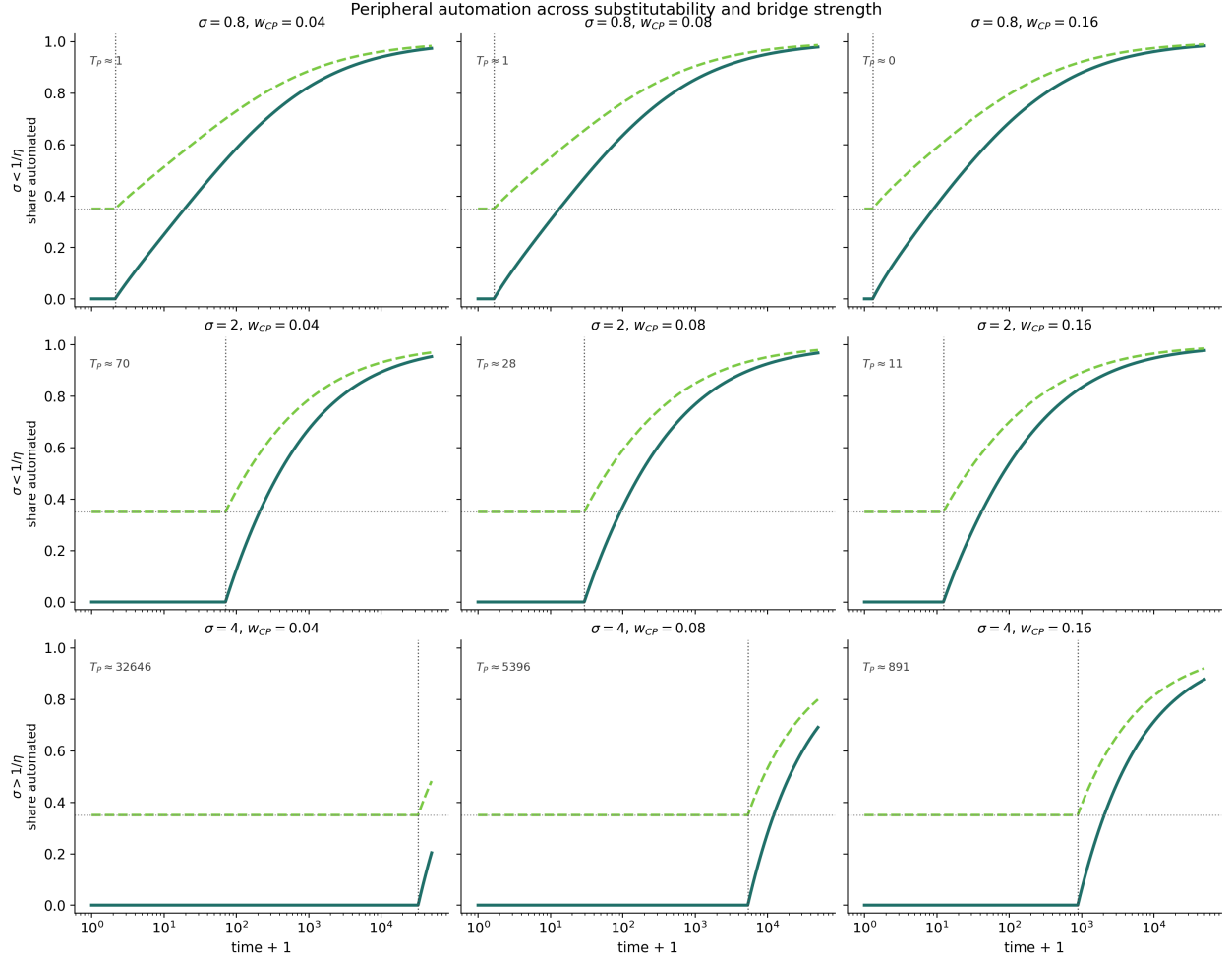


Figure 17

D.3 Capital Accumulation à la Solow

Below we collect additional simulation results with capital accumulation. All parameters are the same as in the baseline, except we assume an exogenous savings rate of $s = 0.02$ and $\delta = 0.01$.

Figure 18: $\sigma = 0.5$ with Solow-style capital accumulation.

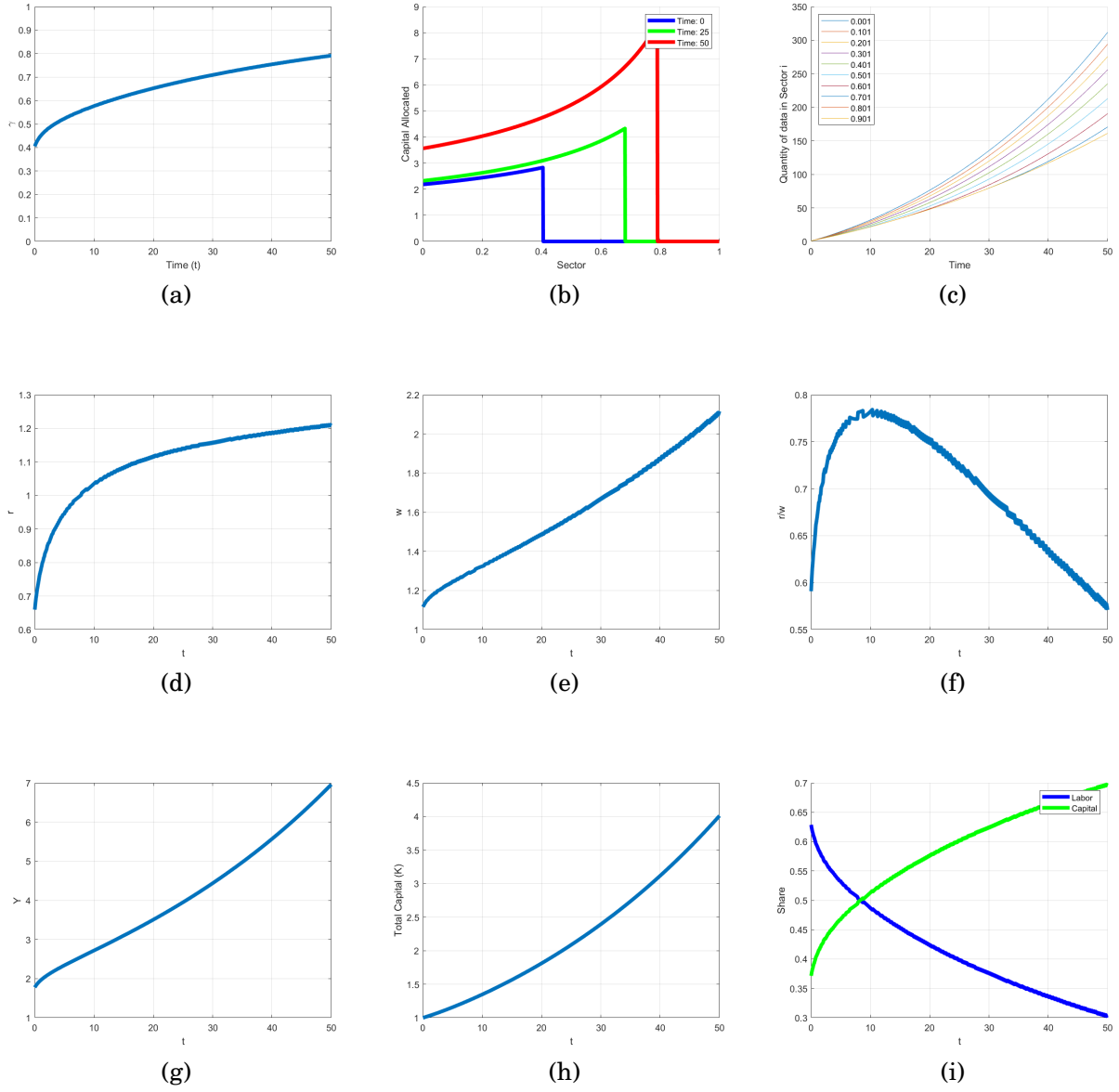
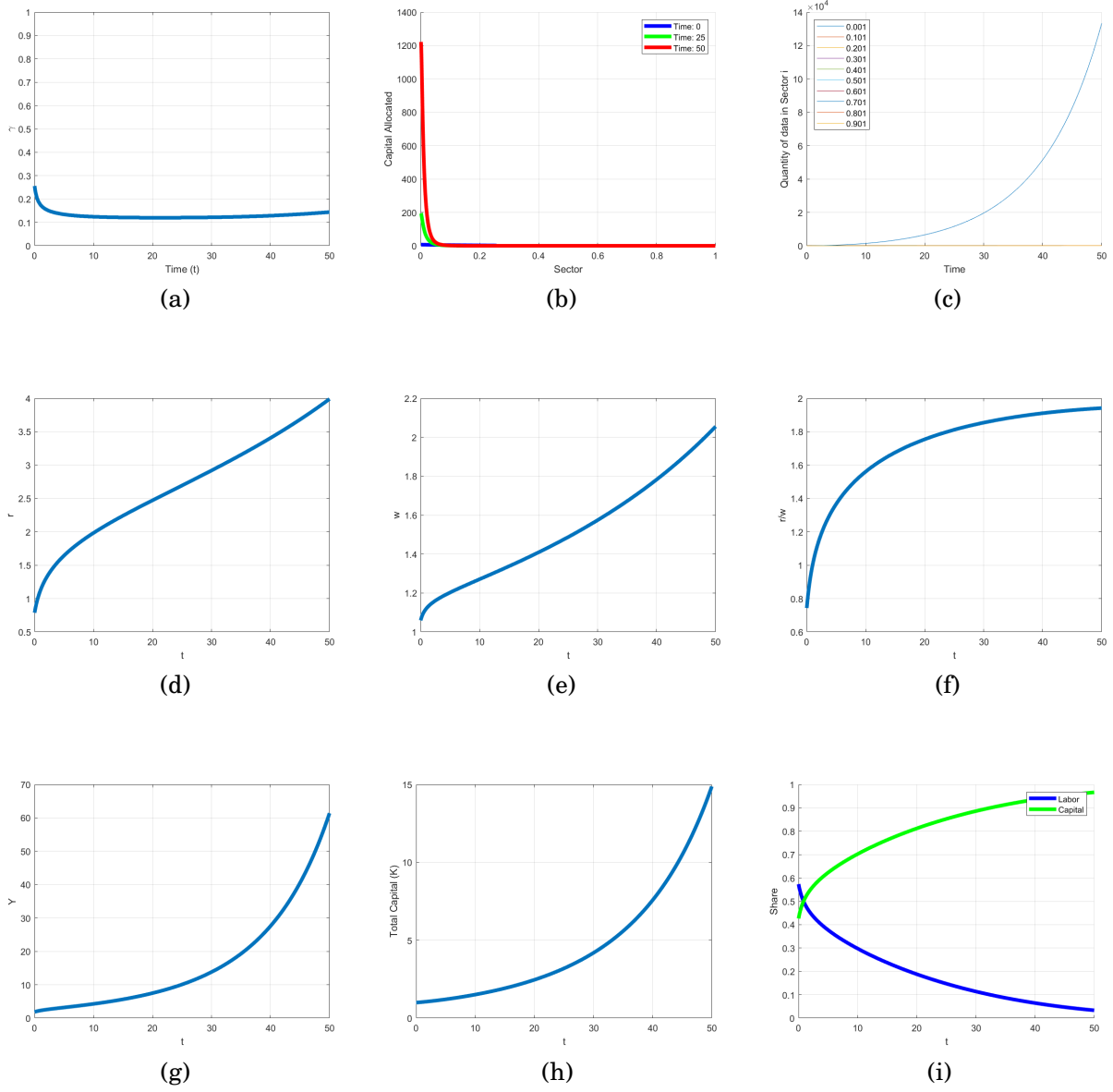


Figure 19: $\sigma = 5.5$ with Solow-style capital accumulation.



D.4 Planner's Optimal Allocation

Below we collect simulation results from the two-task economy. The first two figures are optimal paths chosen by the planner under the ‘AK’ simplification. We set $f_l = 1$, $D_{h0} = 10$, $D_{l0} = 1$, $\eta = 0.2$, $K = 1$, and $\rho = 0.05$.

Figure 20: Planner's allocation at $\sigma = 0.5$

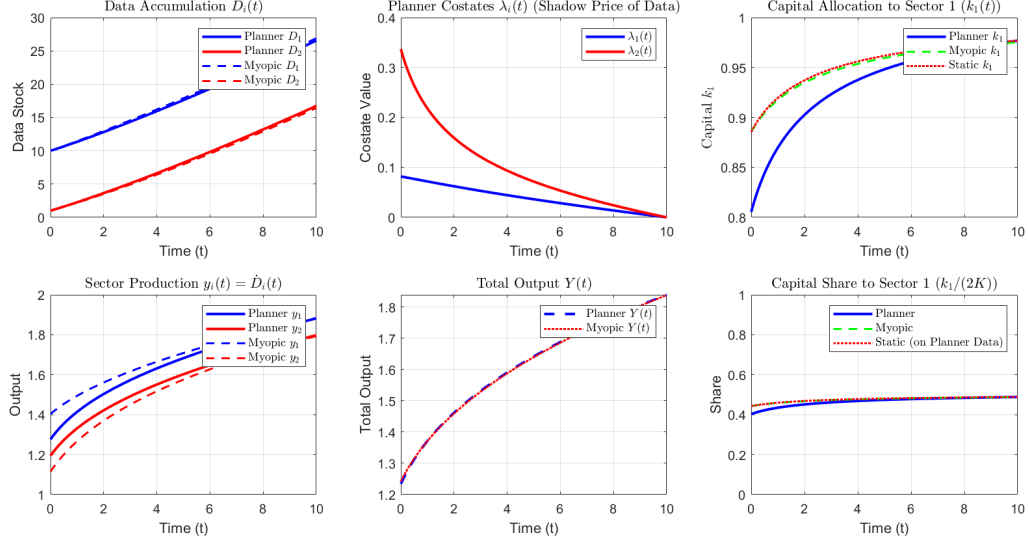
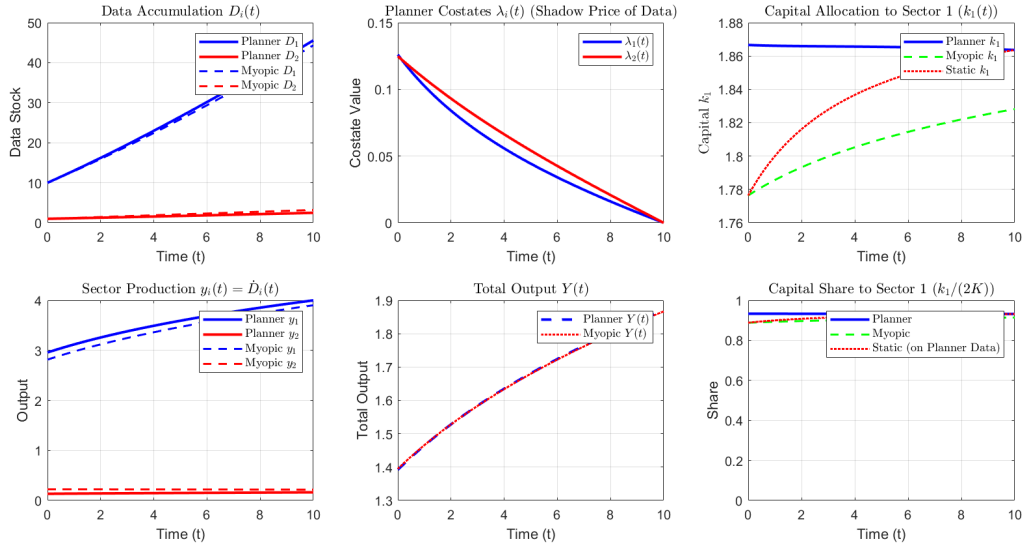


Figure 21: Planner's allocation at $\sigma = 5.5$



Next, we simulate the complement case with labor included. We set $\sigma = 0.5$, $f_h = f_l = 1$, $D_{h0} = 100$, $D_{l0} = 1$, $\eta = 0.2$, $K = L = 1$, and $\rho = 0.05$.

Figure 22: Planner's allocation at $\sigma = 0.5$ with labor included

