Automation Experiments and Inequality

Seth Gordon Benzell
Chapman, MIT, & Stanford

Kyle R. Myers
Harvard & NBER

January 3, 2026

Many experiments study the productivity effects of automation technologies such as generative algorithms. A key test in these experiments relates to inequality: does the technology increase output more for high- or low-skill workers? However, the theoretical content of this empirical test has been unclear. Here, we formalize a theory that describes the experimental effect of automation technologies on worker-level output and, therefore, inequality. Worker-level output depends on a task-level production function, and workers are heterogeneous in their task-level skills. Workers perform a task themselves or delegate it to the automation technology. The inequality effect of improved automation depends on the interaction of two factors: (*i*) the correlation in task-level skills across workers, and (*ii*) workers' skills relative to the technology's effective skill. In many cases we study, the inequality effect is non-monotonic — as technologies improve, inequality decreases then increases. The model and descriptive statistics of skill correlations generally suggest that the diversity of automation technologies will play an important role in the evolution of inequality.

Benzell: benzell@chapman.edu. Myers: kmyers@hbs.edu. This project received support from the Harvard Business School. There are no financial relationships or other potential conflicts of interest that apply to any of the authors. We appreciate helpful feedback from Rem Koning, Alex Moehring, and Philip Trammell.

1. Introduction

Recent advances in the capabilities of algorithms have led to a surge of investigations into the economics of automation technologies. An important, open question is the effect of these technologies on earnings inequality. Researchers have approached this question using both macroeconomic and microeconomic methods. In this paper, we are interested in what can and cannot be learned from the microeconomics.[1]

Specifically, we focus on the microeconomics of automation experiments: studies that use exogenous variation in the availability of a technology to estimate productivity effects. These experiments are valuable because they yield internally valid estimates (typically, in partial equilibrium) and provide managers and policymakers information about technologies at a relatively high frequency. In order to estimate inequality effects, experimenters often test whether the technology increases the output of higher- or lower-skilled workers with the same job.[2] In June of 2025, *The Economist* summarized the inequality effects from the automation experiments to date: "*Although early studies suggested that lower performers could benefit ... newer studies look at more complex tasks... In these contexts, high performers benefit far more than their lower-performing peers,*" which implies that the skill *level* of the workforce is an important feature and suggests a future of skill-biased technical change, as implied by the article's title "*How AI will Divide the Best from the Rest,*" (The Economist 2025).

But what exactly do we learn if we see an automation experiment lead to more, or less, inequality? What features of the workforce are relevant to the inequality effect? If a technology increases inequality now, will it continue to increase inequality as it becomes more capable?

To answer these sort of questions, we introduce a model in the task-based tradition; output is increasing in task-level inputs. However, unlike many macroeconomic models of task-based production, we explicitly model jobs as production functions at the *individual level*. Individual workers are defined by their endowments of skill in each task, which motivates our focus on a feature of the workforce that is rarely highlighted: the correlation of workers' skills across tasks. This correlation describes, for example, whether the most analytical physicians are also the most empathetic, or whether the best fielder on the baseball team is also the best batter.

We are focused on the context of short-term automation experiments, in which the

[1]The macroeconomic work typically asks: after general equilibrium effects play out, will the new distribution of jobs and earnings be less or more unequal? The answer appears often is that technologies that reduce demand for certain types of labor in the short run may increase demand for that labor in the long run, and vice-versa. For overviews covering historical and current waves of automation, see: Katz and Autor (1999); Acemoglu and Autor (2011); Acemoglu and Restrepo (2019).

[2]For examples of automation trials that report inequality results, see: Noy and Zhang (2023); Chen and Chan (2024); Choi, Monahan, and Schwarcz (2024); Kreitmeir and Raschky (2024); Cui et al. (2025); Brynjolfsson, Li, and Raymond (2025); Dell'Acqua et al. (2025); Kanazawa et al. (2025); Kim et al. (2025); Otis et al. (2025); Roldán-Monés (2025).

sample population is held fixed. The model allows for either either positive or negatively correlated skills among workers. Automation technologies can perform one (or more) of the tasks in production, and a technology's capability may be below or above a given workers' skill. Workers are assumed to have rational expectations over whether using a technology will improve their performance.[3]

The model provides clear predictions about the change in inequality (i.e., the difference in output between worker types) given a marginal increase in an automation technology's capability. Even in simple settings with only two tasks, perfect skill correlations, and single-task automaton, we obtain results that may seem counter-intuitive at first, but become very clear through the lens of the model. Changes in inequality do not depend on the absolute level of workers skill, as the *The Economist* quote above implies or as intuition might suggest. It is the interaction of skill correlation and technological capability that determines the inequality effect.

For example, consider two samples of workers with identical, two-task CES production functions and identical task-level skill distributions. In one sample, skills are positively correlated across workers and in the other they are negatively correlated. Here, an increase in the capability of a technology that automates one task can *decrease* inequality in one sample while *increasing* inequality in the other.

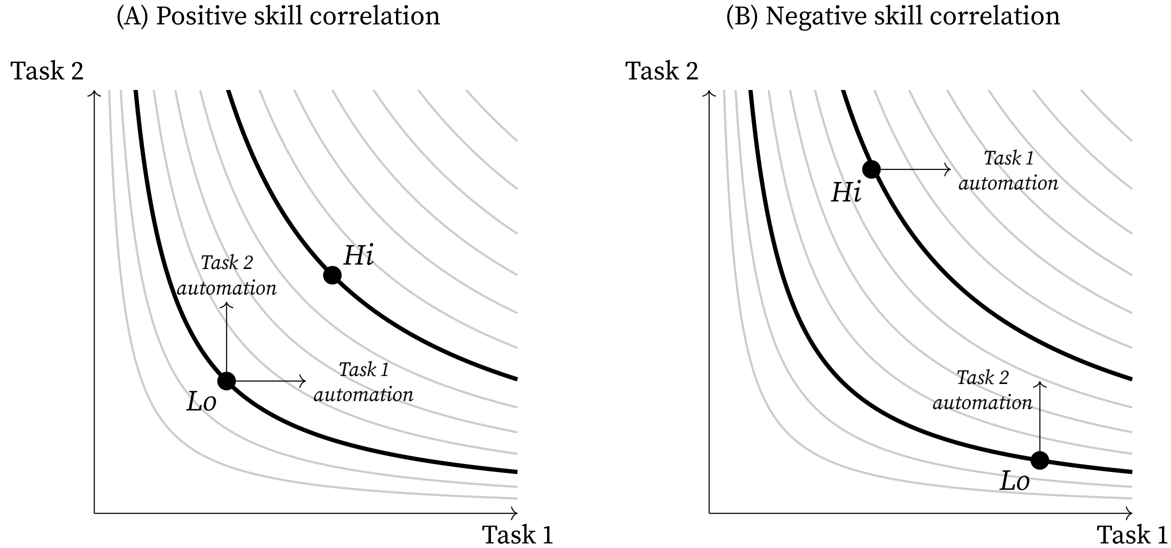
Figure 1 illustrates this case. In Panel A, skills are positively correlated and in Panel B skills are negatively correlated. However, the cases are otherwise identical: the same distribution of total productivity across workers, and the same distribution of individual abilities. Still, as the figure illustrates, a marginal increase in automation technology will have very different effects on inequality across the pair of cases.

In the positive skill correlation case, as automation technology improves from an initial capability of zero, the first user will be the low type. However, in the negative correlation case either type may use the technology first – in this example, the high-type benefits first from task 1 technology improvements. Thus, at least initially, the low type will benefit first and inequality will decrease in the Panel A case. But the high type will benefit first and inequality will increase in the Panel B case.

A second result is that the inequality effect need not be monotonic in the automation technology's capability. Continuing with the example in Figure 1 Panel A, let's again focus on Task 1 automation. As just noted, the low type will be the initial users, which decreases inequality initially. However, once the technology's capability surpasses the high type's skill and the high type begin usage, output differences will be governed by differences in Task 2 skills. Now, since the high type is more skilled at Task 2 (and tasks are complementary),

[3] This assumption is not innocuous, as some automation experiments do find workers to perform worse than placebo when given access to generative algorithms (Dell'Acqua et al. 2025; Otis et al. 2025), which suggests the presence of biased beliefs.

FIGURE 1
Initial Changes in Output after Automation



Note: Plots isoquants of production possibilities, highlighting the isoquant for a worker with higher output (*Hi*) and a worker with lower output (*Lo*) in two scenarios, where task-level skills are positively correlated (Panel A) or negatively correlated (Panel B). Workers pre-automation endowments are shown by the black circles. If a task-specific automation technology is made available and the technology's capability exceeds the workers skill endowment, they will use the technology, which will shift them to a new production frontier as shown by the arrows. For illustrative purposes, the first worker to benefit from automation of a given task is highlighted.

inequality will *increase* as the Task 1 automation technology improves.

Allowing for more than one task to be automated yields additional results. One is that the shortest path to equality is when automation improvements are balanced across tasks. This motivates the potential importance of a diversity of technological advances when it comes to the pursuit of equality. But importantly, "equality" in our model is only possible due to two key assumptions: (*i*) workers are homogeneous with respect to the non-automatable skills; and (*ii*) we allow for the possibility that technologies' capabilities surpass all workers. The reality of these assumptions (or the timeline on which they might occur) is debatable, so automation may ultimately still lead be inequality in practice.

While several dynamics are possible, ours is not an "anything goes" model. When information is available about workers' task-level skills and the relative capability of the automation technology, our model can be use for predicting or extrapolating inequality effects. It is designed to be useful for researchers, organizations, managers and policymakers as they design and interpret automation experiments.

While we focus on partial equilibrium and microeconomics, our model suggests new insights into the macroeconomics of automation. In particular, the model's focus on skill correlations emphasizes the importance of understanding whether the "experiment" is occurring at the job-, organization-, sectoral-, or societal-level.

Existing theory and data provide some guidance on when and where we should expect skill correlations to be positive or negative. We discuss this in the latter portion of the paper and report some country-level skill correlations (based on SAT scores) from the US National Longitudinal Survey of Youth (BLS (2024)). While cognitive skills are consistently positively correlated at the social level, cognitive and physical skills are often negatively correlated. We also apply Kremer's (1993) O-ring model of production to show why skills may be negatively correlated within firms even if they are positively correlated in the population.

After a brief literature review below, the paper is organized as follows: Section 2 lays out the model in detail; Section 3 derives results from the model; Section 4 connects our theoretical results to evidence on skill correlations across the economy and how the level of aggregation (e.g., firms versus populations) has an effect on these correlations; and Section 5 concludes with a discussion of our model's implications, limitations, and possible extensions.

Related Literature

Our paper builds on research on the productivity impact of emerging technologies, most notably, generative algorithms, theories that model jobs as multi-task production functions, and the broader economics of automation.

Generative Algorithms and Inequality. Table 1 summarizes recent automation experiments that have estimated inequality effects due to generative algorithms. As the table shows, findings are decidedly mixed.

Our theoretical framework helps rationalize these seemingly inconsistent findings. Given the rapid pace of algorithmic progress, experimental results at one point in time may not predict effects as technologies continue to advance. Moving forward, experiments that measure not only productivity effects but also: task-level skill distributions among the sample; which types use the technology most intensively; and the technology's capability relative to worker skills will provide more robust guidance for anticipating distributional consequences.

Jobs as Production Functions. We are not the first to treat jobs as production functions. One close example is Autor and Handel (2013; see their Eq. 2), which models job-level output as an exponential function of workers' task-level skills and job-task-specific elasticities. There, the focus is not on automation, but rather on predicting job-level sorting. Deming (2017) measures the relationship of workers' math and social skills to wages, finding that they are complementary in production. But he does not focus on the implications of different correlations in these skills across workers. Dessein and Santos (2006) also consider jobs

TABLE 1
Inequality Results in Recent Automation Experiments

	job	*methodology*	*inequality effect*
Brynjolfsson, Li, and Raymond (2025)	Customer support	Natural Experiment	↓
Chen and Chan (2024)	Ad copywriting	RCT	↓
Choi, Monahan, and Schwarcz (2024)	Legal work	RCT	↓
Cui et al. (2025)	Coding	RCT	↓
Dell'Acqua et al. (2025)	Consulting	RCT	↓
Kanazawa et al. (2025)	Taxi driving	Natural Experiment	↓
Maršál and Perkowski (2025)	Banking	RCT	↓
Noy and Zhang (2023)	Professional writing	RCT	↓
Kim et al. (2025)	Investment analysis	RCT	↑
Otis et al. (2025)	Entrepreneurship	RCT	↑
Roldán-Monés (2025)	Debating	RCT	↑

Notes: Summarizes recent automation experiments involving generative algorithms. Studies are sorted based on whether they find automation to decrease (↓) or increase (↑) inequality, whether they use a formal randomized controlled trial (RCT) or a natural experiment, and then alphabetically.

as bundles of tasks, although the focal question is how organizations endogenously decide which tasks are bundled together. Our model's highlighting of task-level correlation in skills should prove useful in future work using worker-level production functions, specifically, and the division of labor more generally (e.g., Becker and Murphy 1992; Autor, Levy, and Murnane 2003). We review empirical studies of skill correlations later in Section 4.1.

Automation in General. Related work in the broader economics of automation includes Athey, Bryan, and Gans (2020), who develop a framework for optimal delegation between humans and algorithms, and Xu et al. (2025), who study how generative algorithms affects organizational structure and the allocation of tasks within firms. Trammell (2025) focuses on the automation of tasks when there is learning-by-doing in a sequence of tasks that constitute the workflow for a job. There is also an emerging stream focused on the design of human-technology collaborations (e.g., Kleinberg et al. 2018; Mullainathan and Obermeyer 2022; Agarwal et al. 2024; Agarwal, Moehring, and Wolitzky 2025). Regarding the macroeconomics of automation and inequality, Katz and Autor (1999), Acemoglu and Autor (2011), and Acemoglu and Restrepo (2019) provide comprehensive overviews of how technological change affects the wage structure through general equilibrium channels. More recently, Autor and Thompson (2025) examine how generative algorithms affect the returns to expertise, while Garicano and Rayo (2025) studies implications for training and apprenticeship.

Our partial equilibrium approach abstracts from general equilibrium effects in order to isolate the direct productivity channel, which yields predictions about which workers benefit from automation in the short run before labor markets adjust. This is the relevant experimental context and the focus of this paper. However, a natural extension of our model is to embed it into leading general equilibrium models of automation (e.g., Zeira 1998; Acemoglu and Restrepo 2018) to predict longer-term and general equilibrium effects.

2. Model

Here, we outline our model of jobs as production functions and the automation of tasks. The tasks of the production functions and the skills of workers are exogenous and fixed. This short-run, partial-equilibrium view reflects the format of most automation experiments, given their pursuit of internal validity.

2.1. Jobs as Production Functions

There is a single job where each worker i produces a non-negative quantity of some output Y. The job involves a set of tasks $t = \{1, 2, ..., T\}$, and workers' output is a function of their vector of skills in each task. Implicitly, workers inelastically supply a unit of effort towards each task. In the Appendix, we allow for workers to endogenously supply their labor such that time-on-task depends on both their skill levels and the automation technology's capability. While this complicates the analysis, it does not impact our key results.[4] Workers are differentiated by their task-specific skills $S_{it} > 0$. Since labor is inelastic, we can describe each worker's exogenous (job-specific) skill endowment across tasks as a vector $\mathbf{S}_i = (S_{i1}, ..., S_{iT})$.

Output is produced according to a general CES production function:

$$Y_i = \Big(\sum_t \alpha_t S_{it}^{\rho}\Big)^{1/\rho} \qquad \text{where} \qquad \rho = \frac{\sigma - 1}{\sigma} \leq 1\,,\ \sigma \geq 0\,,\ \sum_t \alpha_t = 1\,, \tag{1}$$

where σ is the job-specific elasticity of substitution, and α_t are the task-specific input shares.

While most jobs contain more than two tasks, in the context of automation experiments, it is useful to partition all tasks into two groups: one set that may jointly delegated to the new technology, and another set that may not be delegated. Therefore, we simplify our production function to two tasks ($T = 2$). Skill levels are normalized so each task has equal input shares

[4] In other words, in the baseline model, output is a function of the worker's skill along each dimension, not any endogenous effort decisions. Still, the result of time-on-task being stable will be chosen to be fixed endogenously in a Cobb-Douglas setting if technology is interpreted as multiplying rather than substituting for worker skill. We thank Phil Tramell for this last point.

($\alpha_t = 1/2$). This simplifies the production function to:

$$Y_i = \left(\frac{1}{2}\left(S_{i1}^{\rho} + S_{i2}^{\rho}\right)\right)^{1/\rho} . \tag{2}$$

For some results, we focus on the edge case of $\rho \to 0$, which yields the familiar Cobb-Douglas production function:

$$Y_i = (S_{i1} \times S_{i2})^{1/2} . \tag{3}$$

In Appendix A, we consider the case where workers' labor supply is not fixed. Specifically, we focus on the Cobb-Douglas case and introduce a convex cost of effort that pins down workers' optimal choices given their skills and, in the case of an automation experiment, the capability of the technology. This of course changes the specific values involved, but all of the qualitative results we derive below persist. Thus, endogenizing workers' effort choices yields no significant insights beyond the simpler case of a fixed, inelastic labor supply.[5]

2.2. Automation Experiments

We model automation experiments as the exogenous offering of a technology that can automatically perform a task. We denote the capability of the automating technology for task t as $A_t \geq 0$. As an input, the automating technology may be worse, equal to, or better than labor (i.e., $A_t \lesseqgtr S_{it}$). We assume that workers have rational expectations about the technology's capability; thus, i will rationally choose to use it if $A_t > S_{it}$ and not use it otherwise.

The production function with single-task automation becomes:

$$Y_i = \left(\frac{1}{2}\left(\max(S_{it}, A_t)^{\rho} + S_{it^-}^{\rho}\right)\right)^{1/\rho} . \tag{4}$$

where t is the task being automated and t^- is the other task. When we consider cases where both tasks are being automated, we will use the max function for both tasks.

2.3. Skill Correlations

We consider the two limit cases for the skill correlation structure, both in which there is ex-ante (pre-experiment) differences in output levels. There are two groups of workers: those with initially high output ($i = Hi$) and those with initially low output ($i = Lo$).[6] Table 2 summarizes our negative and positive correlation cases, noting that we use the variables C

[5]Future work that endogenizes the *creation* of new tasks by the worker may prove fruitful.

[6]This binary categorization is simple, and it reflects the common empirical exercise of proxying for workers' skills using their pre-experiment output levels to divide them into two groups (e.g., Noy and Zhang 2023; Chen and Chan 2024; Choi, Monahan, and Schwarcz 2024; Cui et al. 2025; Brynjolfsson, Li, and Raymond 2025; Dell'Acqua et al. 2025; Kanazawa et al. 2025; Kim et al. 2025; Otis et al. 2025; Roldán-Monés 2025).

and B to describe the workers comparative and/or absolute advantages, where $1 < C < B$.[7]

TABLE 2
Skill Correlation Scenarios — Worker (i) and Job Task (t)

	positive corr.		*negative corr.*	
	$t = 1$	$t = 2$	$t = 1$	$t = 2$
$i = Hi$	B	BC	B	1
$i = Lo$	1	C	1	C

Notes: Reports the values of workers' task-specific skills for the positive and negative correlation cases; $1 < C < B$. Therefore, without loss of generality, in the negative correlation case skill at task 1 is more dispersed.

2.4. Inequality Measures and Effects

There are many ways to measure inequality. In practice, one of the most common ways worker-level output inequality is evaluated is simply the first difference in output levels between workers. The change in inequality is estimated by testing whether workers with high initial output levels (pre-automation) benefit more or less than workers with low initial output levels. If those initially high-output workers benefit more, inequality is said to have increased, and vice versa. Thus, we define Δ as:

$$\Delta \equiv Y_{Hi}(A_t) - Y_{Lo}(A_t) ,$$

which provides our "first-difference" measure of inequality. Below, we'll also investigate the absolute value of Δ as a measure of inequality that is agnostic as to the source of the differences. This leads us to our main (partial equilibrium) effect of interest:

$$\partial \Delta / \partial A_t ,$$

which describes how the output gap between workers of groups H and L changes as automation technologies are introduced and/or improved.[8]

As we discussed in the introduction, many automation experiments motivate a focus on $\partial \Delta / \partial A_t$ with concerns about long-run general equilibrium effects. This is clearly heroic in the absence of a structural model. Estimating the general equilibrium impact of a technology on inequality requires understanding not only the direct productivity effect of a technology on

[7]Note this analytical setup differs slightly from the data-generating process used to produce Figure 1. There, the focus was on ease of visual understand, and here the focus is on analytical tractability.

[8]Note that we may normalize the initial productivity of either group to 1, making this linear difference also interpretable as a percentage difference.

worker output, but also elasticities of substitution across types of worker output, and many other elasticities (e.g., workers changing jobs, or employers changing the task composition of jobs).

More modestly, one might hope to naively extrapolate from the context of one experiment to another related one: For example, if a given technology reduces productivity inequality in a given firm, it might be expected to have a monotonic effect as the inequality increases. Equally plausibly, it might have a similarly signed (short term, partial equilibrium) effect when deployed in a new sample population.

Formally, the first inference could be stated as an assumption that $\partial\Delta/\partial A_t$ is constant, or at least the same sign, as A_t increases or samples change. We investigate the first imputation by solving for the second derivative, $\partial^2\Delta/\partial A^2$.

To further explore heterogeneity, we also solve for the ways in which the inequality effect depends on: (*i*) how large the absolute advantage of the high-type workers is, $\partial^2\Delta/\partial A\partial(B/C)$; and (*ii*) how substitutable the two tasks are in the production function, $\partial^2\Delta/\partial A\partial\rho$.

Furthermore, we'll investigate the case where both tasks may be automated and illustrate how the effect of inequality jointly depends on the capabilities of the two technologies, $\partial^2\Delta/\partial A_1\partial A_2$.

Sometimes in our model a technology can be powerful enough to flip the initially lower-type worker to being more productive than the high-type worker. This is possible in the negative correlation case, when task one (the task the high-type is initially better at) reaches a high level of automation quality. Because of this possibility, we also focus our attention on the absolute value of Δ:

$$|\Delta| \equiv |Y_{Hi}(A_t) - Y_{Lo}(A_t)| \ .$$

Across the literature, other common metrics of inequality include statistics such as the coefficient of variation or the Gini coefficient. These metrics typically involve a scaling of absolute differences by some measure of aggregate levels.[9] This scaling is useful for making comparisons across contexts, since it converts absolute values into relative magnitudes. However, since we will always be considering the same context in theory, we will rely on our simpler measures.

[9] For example, in our setting, the coefficient of variation is simply: $|Y_{Hi} - Y_{Lo}|/(Y_{Hi} + Y_{Lo})$; and the Gini coefficient is simply the coefficient of variation divided by 2.

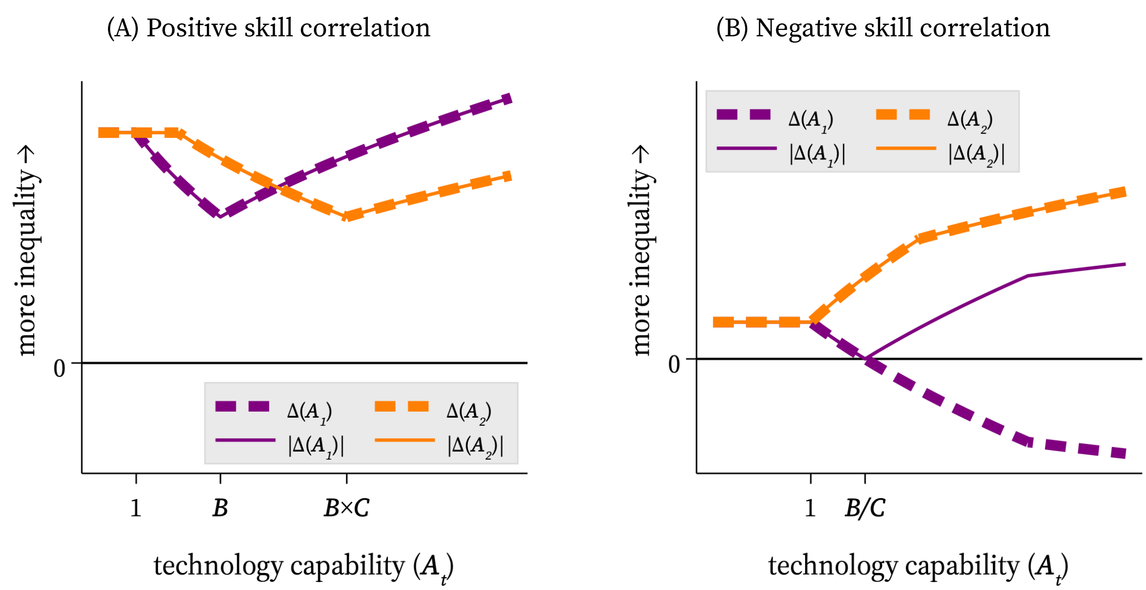
3. Model Results

3.1. Single-task Automation

3.1.1. Cobb-Douglas Case

We begin with a focus on the simple case of Cobb-Douglas production and single-task automation. Figure 2 reports the results visually. It visualizes inequality as a function of the correlation in workers' skills (per panel), which task is automated (per line color), and the capability of the automation technology (per the x-axis).

FIGURE 2
Inequality and Single-task Automation — Cobb-Douglas Production



Note: Plots the level of inequality across the two groups of workers assuming Cobb-Douglas production per first-difference inequality (Δ; solid line) or absolute inequality ($|\Delta|$; dashed line) depending on skill correlations, the task being automated, and the capability of the technology.

A first result the figure illustrates is that when workers' skills are positively correlated, Δ must be weakly greater than zero. Equivalently, Δ and $|\Delta|$ align. In other words, when skills are positively correlated, there is no possibility for technology to "boost" the low-types above the high types. This intuitive result holds in the general CES case.

On the other hand, Figure 2 Panel A also shows that, regardless of whether task 1 or 2 is automated, there will be a non-monotonic relationship between inequality and the automation technology's capability. First inequality declines, then it increases. Initially, when first introduced, automation helps the low type "catch-up" to the high-output type. This is because the technology's capability is greater than the low type (so, the low type uses the tool), but lower than the high type (so, the high type relies on their own skill). But later, the technology is better than both types, and the high-type's absolute advantage in the non-automated task becomes the key determinant of output differences and inequality grows due to the complementarity in the tasks. In short, when skills are positively correlated, low-capability partial automation helps the low-types to catch up; but high-capability automation allows the high-type to shine.

In the case of negatively correlated skills, visualized in Panel B, the results are strikingly different. Here, the effect of single-task automation *is* monotonic in *first-difference* inequality for both tasks. When the task the high-skill are better at is automated, the measure decreases. But when the task the low-skill is better at is automated, the high-skill pull farther away.

However, in the special case of negatively correlated skills where the skill the high-type is better at is being automated, the reduction in Δ is so extreme, the measure can turn negative. In this case, the effect on absolute inequality is non-monotonic: a certain point, more better automation of the skill that the low-types are worse at increases their lead over the ex-ante high-types (see Figure 2B). This divergence of first-difference and absolute inequality highlights the importance of tracking workers over time. Cross-sectional analyses of automation experiments may reveal no change in inequality while masking a major re-ordering of workers.

The specific points at which these non-monotonic switches in inequality occur will be context-specific. But overall, the simple Cobb–Douglas case clearly illustrates how improvements in automation may increase or decrease output inequality depending on the workers' skill correlations and the capability of the technology relative to each workers' skills. Furthermore, as we show in Appendix A, this non-monotonicity persists in a model where task-level labor supply is endogenous.

3.1.2. General CES Results

Table 3 generalizes these results to the CES production function. The table reports the signs of first and second derivatives of absolute inequality with respect to automation capability, providing a more complete picture of when and why automation's inequality effects may reverse.

In all cases, the sign of this second derivative with respect to the technology's capability ($\partial^2|\Delta|/\partial A^2$) is opposite to the sign of the first derivative, which stems from the concavity of the production function.[10] This matters for practice: a firm piloting automation today may observe inequality changing, and even before the dramatic non-monotonicity occurs, the rate of change in inequality will decline.

[10]Except at kinks (i.e. the quality level where a technology first becomes worth using for a group) where the second derivative is undefined.

TABLE 3
Absolute Inequality and Automation

technology	*tech. capability*	*sign* $\frac{\partial\lvert\Delta\rvert}{\partial A}$	*sign* $\frac{\partial^2\lvert\Delta\rvert}{\partial A^2}$	*sign* $\frac{\partial^2\lvert\Delta\rvert}{\partial A\partial B/C}$	*sign* $\frac{\partial^2\lvert\Delta\rvert}{\partial A\partial\rho}$
Panel (A): Negative skill correlation					
A_1	$1 < A_1 < A^*$	$(-)$	$(+)$	0	$(+)$
A_1	$A^* < A_1$	$(+)$	$(-)$	0	$(-)$
A_2	$1 < A_2 < C$	$(+)$	$(-)$	$(+)$	$(-)$
A_2	$C < A_2$	$(+)$	$(-)$	0	$(-)$
Panel (B): Positive skill correlation					
A_1	$1 < A_1 < B$	$(-)$	$(+)$	0	$(\pm)^a$
A_1	$B < A_1$	$(+)$	$(-)$	$(+)$	$(-)$
A_2	$C < A_2 < BC$	$(-)$	$(+)$	0	$(-)$
A_2	$BC < A_2$	$(+)$	$(-)$	$(+)$	$(-)$

Notes: Reports the sign of the first and second derivatives of the absolute inequality effect ($\partial|\Delta|$) for the general CES production function with automation depending on whether workers' skills are negatively correlated (Panel a) or positively correlated (Panel b). $A^* = (B^\rho + 1 - C^\rho)^{1/\rho}$. $(\pm)^a$: sign is positive if $A < C$ and negative if $C < A < B$.

The cross-derivatives with respect to the high-type's skill advantage ($\partial^2|\Delta|/\partial A\partial B/C$) reveal a few instances where initial skill gaps can amplify automation's inequality effects. In the negative skill correlation case, the skill gap (B/C) plays no role in the inequality effect of automation. However, this derivative is positive in the positive correlation case once technology's capabilities exceeds both workers' skills — a larger pre-existing skill gap means automation will increase inequality more sharply. For organizations considering automation, this implies that the same technology will have different distributional consequences depending on how heterogeneous the workforce is to begin with. A narrow skill distribution may mute inequality changes; a wide distribution may amplify them.

The cross-derivatives with respect to task complementarity ($\partial^2|\Delta|/\partial A\partial\rho$) indicate whether automation's effect on inequality is amplified or not in jobs with more substitutable tasks ($\rho \to 1$). In the case of negative skill correlation, task substitutability behaves just as the second derivative; more substitutability mutes the inequality effect. With positively correlated skills, the cross derivative is negative for most technological capabilities, which amplifies inequality effects at low capability levels but depresses them at higher capability levels.

Practically, these comparative statics tell us that predicting automation's inequality effects requires knowing not just current worker skills and technology capabilities, but also the job's production structure and the trajectory of technological improvement. Perhaps most strikingly, the across all cases the comparative static of improving technology's capabilities for a single task at a time all reveals that the effect of sufficiently super-human ability at a given task (i.e., $A_t \to \infty$) *must* be to increase inequality.

3.2. Full Automation

Returning to the Cobb-Douglas formulation for simplicity, Figure 3 extends the analysis to consider automation of both tasks, revealing how the task *composition* of automation technologies shapes inequality outcomes. Each panel plots absolute inequality compared to the pre-automation period as a function of both technologies' capabilities. The contrast between panels highlights how skill correlations determine not just whether inequality rises or falls, but which combinations of technologies prove most equalizing.

In both cases, there is convergence to zero inequality when automation is sufficiently advanced; however, we note again the fact that our model assumes workers are homogeneous with respect to non-automatable tasks (since there are none in our model). This is an important abstraction. Until the economy achieves full automation, there will be non-automated tasks, so the "equality" we arrive at should be understood as representing the amount of irreducible inequality determined by all non-automatable tasks.[11]

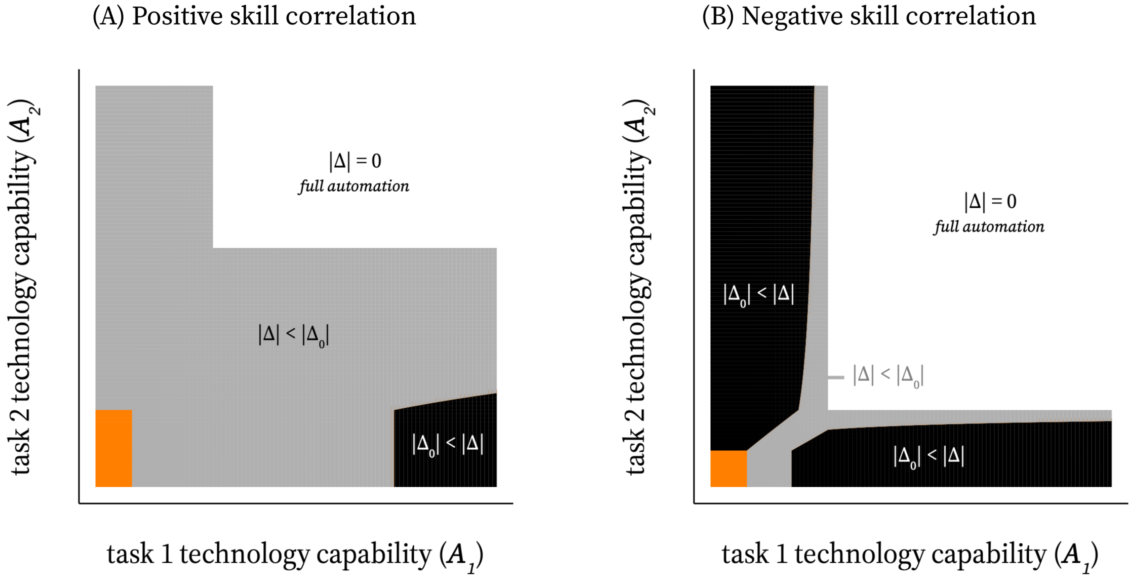
With that caveat in mind, Figure 3 reveals the asymmetric paths to equality. When automation advances differentially across tasks, inequality remains or grows — the technology portfolio is unbalanced relative to workers' skill distributions. This illustrates how the balance of technological capabilities is relevant to inequality. As more tasks are automated, the dimensionality along which workers differ shrinks, and inequality increasingly reflects skill differences in the remaining non-automated tasks.

With positive correlation (see Figure 3A), much of the capability space involves decreases in inequality compared to pre-automation periods. Effectively, technologies are always helping the low type catch up. Only if the task 2 technology has very low capability and task 1 technology has very high capability can inequality exceed the pre-automation period.

With negative correlation (see Figure 3B), there is a much greater opportunity for increases in inequality. The intuition follows from the single-task results: when skills are negatively correlated, different workers are disadvantaged in different tasks, so equalizing outcomes requires helping both types of workers. A technology suite that is very capable in task 1 helps low-types, leaving low-types' comparative advantage in task 2 intact, which remains a (potential) source of inequality.

[11] This might represent the desire for human-in-the-loop involvement for its own sake. Alternatively, if humans are not endowed with any essentially non-automatable skills (i.e. allowing for complete automation of the occupation), inequality will be zero or undefined, depending on precise definition.

FIGURE 3
Absolute Inequality and Multi-task Automation — Cobb-Douglas Production



Note: Plots the level of absolute inequality ($|\Delta|$) compared to the initial, pre-automation level of absolute inequality (defined here as: $|\Delta_0|$) depending on skill correlations and the capability of both automation technologies. The results here assume Cobb-Douglas production, and the orange areas indicate regions where neither technology has been adopted (i.e., $|\Delta| = |\Delta_0|$).

3.2.1. Limits and the Long Run

Automation technologies are anticipated to improve considerably, so a natural case to consider is what happens in the limit as these technologies become arbitrarily powerful. As specified above, our model has clear implications for this limit: if only a portion of a job can be automated, then as the automation becomes more powerful inequality will increase among workers doing that same job. Alternatively, if all parts are automated, inequality will be eliminated.

However, in our model (and the vast majority of models of production functions) there are no bounds on the returns to inputs. That is, output is strictly increasing in task-level inputs for infinite levels of inputs.[12] However, in many practical settings of interest, it may be the case that effective skill at one component of a job may be capped. In practice, the nature of the job may be such that, beyond some maximum input level, output is no longer increasing in one or both inputs.[13] This would overturn our results that single-task automation is ultimately

[12]To be clear, this is the norm in theoretical analyses as it prevents any discontinuities in the production function.

[13]For example, if the worker is a plumber that repairs homes, there may be a practical limit as to how the plumbers' physical skills translate into value for a homeowner — one plumber may be stronger than another, but if a bolt connecting two pipes can only be tightened so much, then additional strength yields no additional value.

inequality-increasing (as shown in Figure 2 and Table 3). For instance, consider a Cobb-Douglas job with two tasks where technology is automating task 2, but there is a cap on the returns to that task:

$$Y_i = \Big(S_{i1}, \max\big(\max(S_{i2}, A_2), \overline{S}_2\big)\Big)^{1/2},$$

where $\overline{S}_2$ indicates the point beyond which additional task-2 skill (from the worker) or capability (from the technology) yields no additional output. In this case, the ultimate sign of the inequality effect as technological capabilities improve is undetermined since it depends on the relative magnitudes of $\overline{S}_2$ and A_2 (compared to workers' skill levels and correlations). For example, in our case of positive skill correlation, if $C < \overline{S}_2 < BC$, then the high type have already hit the cap on returns to task-2 input levels, so improvements in A_2 will only ever reduce inequality. In this example, when $C < A_2 < \overline{S}_2$, improvements in the technology raise the low type's output but leave the high type's output unchanged (inequality is reduced), but then once $\overline{S}_2 \leq A_2$, technology improvements have no effect on inequality.

Another case to consider: it may be that a job's productivity is unbounded in it's arguments, but the technological progress necessary for automation to progress may take some time.[14] However, it is important to point out that we take no stance on *how long* said improvements may take. The long run in our model could be months, years, decades, or centuries – which makes a focus on the non-asymptotics of our model equally or more empirically relevant. Still, our framework emphasizes that it is not the absolute capability of the technology that is relevant, rather it is the technology's capability *relative* to the skill levels of different workers. It is when technologies cross certain thresholds based on workers' skill levels that influence the sign of inequality effects from automation (e.g., the B, $B \times C$ and B/C points flagged in Figure 2).

4. Skill Correlations

The previous section highlighted the importance of understanding skill correlations in the workforce of interest. Here, we review the empirical studies that have estimated skill correlations (Section 4.1), present new evidence on the distribution of skill correlations across sectors (Section 4.2), and show how positive assortative matching in an O-ring model of production (i.e., Kremer 1993) can generate negative skill correlations within firms even when the population exhibits a positive skill correlation (Section 4.3).

[14] To continue the above example, perhaps sufficiently advanced robotic plumbers truly could be "unboundedly" valuable to a homeowner.

4.1. Existing Estimates of Skill Correlations

At least since Willis and Rosen (1979), it has been appreciated that a single index of workers' skill will abstract away from important variation in the labor supply. However, empirical estimates of workers' skills on specific types of tasks remains a challenging pursuit. Approaches to date typically involve accessing confidential data, intense measurement, or the estimation of structural models of labor supply.

Most studies examining multi-dimensional skills have focused on aggregate categories: cognitive skills, social skills, and manual skills. While these categories may be much broader than the connotation of the "tasks" in our model, the logic is the same. Furthermore, these studies generally focus on large populations of workers (e.g., national statistics).

Broadly speaking, among these three categories of skills, cognitive and social skills show the most positive correlation. Estimates of the correlation in these skills spans from approximately 0.1 up to 0.7 (Mayer, Roberts, and Barsade 2008; Baker et al. 2014; Deming 2017; Guvenen et al. 2020; Lise and Postel-Vinay 2020; Girsberger, Koomen, and Krapf 2022; Bárány and Holzheu 2025). Cognitive and manual skills exhibit negative to weakly positive correlations, roughly in the range of –0.4 to 0.1 (Lindenlaub 2017; Lise and Postel-Vinay 2020; Bárány and Holzheu 2025). Social and manual skills consistently show the most negative correlations of these categories, on the scale of –0.4 to –0.6 (Lise and Postel-Vinay 2020; Girsberger, Koomen, and Krapf 2022; Bárány and Holzheu 2025). Within the cateogry of cognitive skills, data persistently reveal a strong positive correlation, for example, of approximately 0.7 across mathematical skills and literacy (or verbal) skills (Hampf, Wiederhold, and Woessmann 2017; Guvenen et al. 2020; Woessmann 2024).

4.2. SAT Score Correlations across the Economy

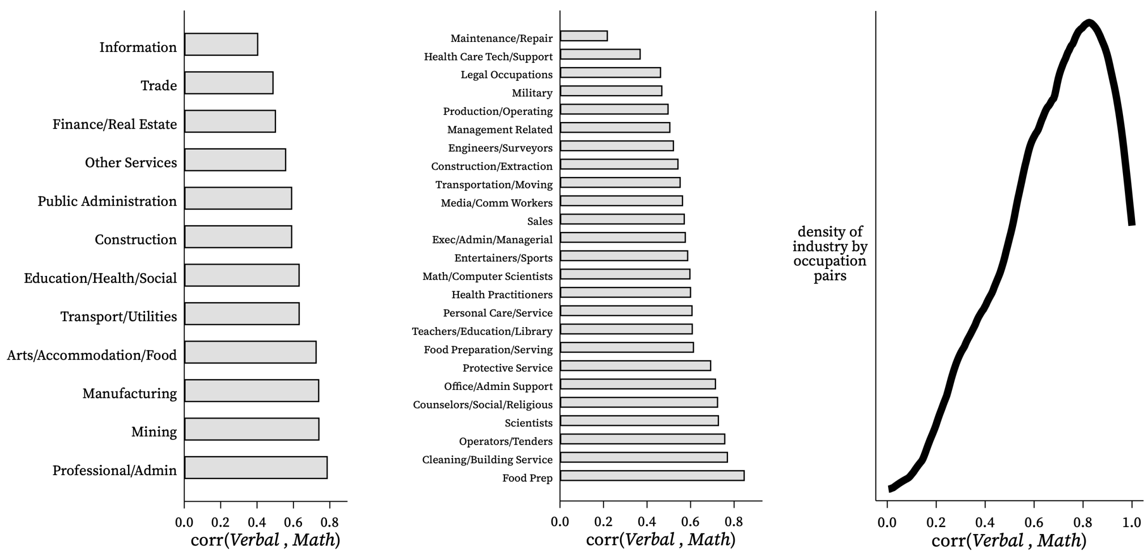
As noted above, prior work has generally found mathematical and verbal skills to be positively correlated per standardized tests. A common source of data used is the National Longitudinal Survey of Youth (NLSY), which reports participants' SAT scores for verbal and math components separately (BLS 2024). Here, we also make use of this data, but report the math-verbal skill correlation separately for alternative industries and occupations. We use the most commonly listed industry and occupation codes for all employed individuals. The public data yield 1,618 observations with both math and verbal SAT scores as well as non-missing industry and occupation codes.[15]

[15] For simplicity and to preserve sample size, we aggregate the industries and occupations into levels slightly broader than listed in the data. SATs are graded out of 800 for each of the two components. For confidentiality, the scores are aggregated into 100 point bins in the public data, which are what we use in this analysis. All correlations are reported per the inclusion of the NLSY97 sample weights, although this makes little practical difference.

Overall, math and verbal SAT scores in our sample exhibit a correlation of 0.61, which is squarely in-line with the estimates summarized in the previous subsection. But our primary goal is to shed some initial light on portions of the economy where skills (at least per this metric) may be more or less correlated.

FIGURE 4
Math and Verbal SAT Score Correlations across the Economy

(A) Across Industries | (B) Across Occupations | (C) Across Industry×Occupations



Note: Plots the correlation of individuals' SAT scores on the verbal and math components from the NLSY97. Panels (A–B) plot average correlations at the industry- (Panel A) or occupation-level (Panel B). Panel (C) plots the distribution of correlations averaged at the industry-occupation-level.

Figure 4 highlights that the strength of the math-verbal score correlation varies significantly across the economy.

Workers in some place exhibit a tight link between these skills, with estimated correlations approaching 0.8 (e.g., professional and administrative services, food preparation, cleaning and building services). While in other places the association is much more modest, on the order of 0.3 to 0.5 (e.g., information sector, legal services). At the granular level of industry–occupation pairs, we find substantial dispersion in math–verbal correlations (see Panel C).

Notably, we do not observe any industry and/or occupation where the correlation is estimated to be negative. Still, our model suggests that the the heterogeneity in skill correlations across the economy would also imply heterogeneous effects of automation on inequality. That is, a single automation technology deployed in these different sectors could yield significantly different changes in inequality.

More speculatively, the uniformly positive correlations help reconcile the fact that, as of this paper, the majority of automation experiments are documenting inequality declines. Presently, many of the technologies in question are both: (*i*) not yet "super-human" in that they are better than the worst workers, but not better than the best workers; and (*ii*) focused on cognitive tasks, in particular, connected to software. Combined with positive (cognitive) skill correlations, our model predicts an inequality reduction in this scenario.

However, the positive math-verbal skill correlations we document here reflect sorting into broad economic sectors and may mask negative correlations at finer levels of analysis within firms or teams where complementary specialization drives task allocation.

4.3. An O-Ring Model with Two Tasks

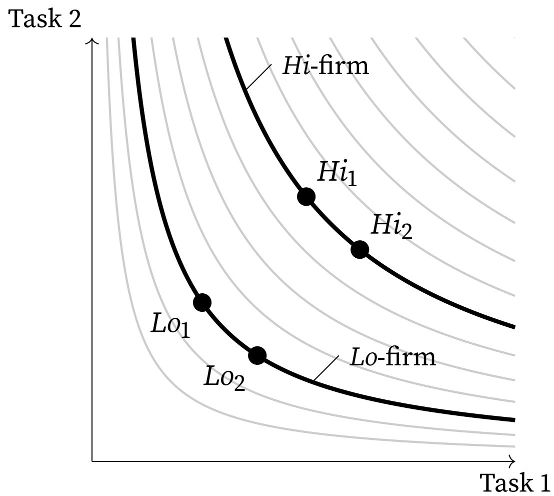
The evidence in the prior section describe skill correlations at the population and occupational levels. However, the focus of many, if not most, automation experiments narrowly focus on a single occupation at a single firm, or possibly even more narrowly defined business units and jobs. In this section, we make the point that the skill correlations within these sample populations may be very different than the population overall.

In order to forecast how worker sorting could lead aggregate skill correlations to diverge from those within an experimental sample, we extend Kremer's (1993) O-ring model. Specifically, we consider the O-ring model when production involves two tasks, unlike the original, single-task model. The model set up is the same as described in Section 2, augmented with additional assumptions to incorporate worker sorting and firm-level production.

The main result of this model is that skills must be negatively correlated within a firm or team no matter the population correlation in these skills. The intuition is straightforward. The O-Ring model assumes that output is complementary in the different inputs to a team, firm, or country. This leads to perfect associative matching in ability across teams or firms: the best teams can't afford any weak links, the least productive firms can't afford higher quality workers. Firms in the middle hire mediocre workers. In our extension, workers' quality is a function of their ability in multiple skills. Because the only ways for a worker to be mediocre overall is to be good at one task but bad at another, or to be mediocre at both, this yields our main finding.

We visualize this intuition in Figure 5. Each productivity isoquant represents both a worker productivity level and a firm type (because all workers at a given firm must have the same productivity level). We plot a pair of high-type workers working at a high-type firm, and a pair of low type workers working at a low-type firm. In the given example, if we assume each node has the same population of workers, the negative correlation in skills within firms occurs despite the fact that there is a positive correlation in skills overall.

FIGURE 5
Skill Correlations and Assortative Matching



Note: Plots isoquants of production possibilities, highlighting the isoquant for workers with higher (*Hi*) or lower output (*Lo*). Workers are chosen to illustrate a scenario where skills are positively correlated in the population (including all four workers). Since there are only two types of workers (and each type exists on the same isoquant), positive assortative matching in the context of two-worker firms will generate a negative within-firm skill correlation.

Generally, the fact that isoquants are weakly downard-sloping requires weakly negative skill correlations in the presence of positive assortative matching.[16]

In the remainder of this section and Appendix B, we formalize this logic.

Consider an economy with a continuum of workers and firms. Each firm employs n workers, and each worker performs two tasks. Worker i has task-specific skill levels $S_{i1}, S_{i2} \in (0, 1]$. Worker i's production function is again the simple Cobb-Douglas form: $Y_i = \sqrt{S_{i1} \times S_{i2}}$.[17] Following Kremer (1993), firm-level output (Z) is the simple product of worker-level output:

$$Z = \prod_{i=1}^{n} Y_i \,, \tag{5}$$

where we have abstracted away from capital used in production for simplicity, and because we are not interested in metrics related to, for example, labor shares.

Skills are distributed jointly log-normal in the aggregate population. Since our focus is on the correlation in skills, we'll focus on the simple case where the means (μ) and standard

[16] We say weakly because, in the case of Leontif production, isoquants have undefined or zero slopes.

[17] We focus on Cobb-Douglas production in this analysis because the general CES form leads to non-linear constraints that do not have straightforward analytical solutions.

deviations (σ) of both skills are equal:

$$\begin{pmatrix} \ln S_{i1} \\ \ln S_{i2} \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu & \sigma^2 & corr_{\text{pop}}\sigma^2 \\ \mu ^{,} & corr_{\text{pop}}\sigma^2 & \sigma^2 \end{pmatrix}, \tag{6}$$

where $corr_{\text{pop}} \in (-1, 1)$ is the population correlation of logged skills.

The multiplicative firm-level structure ensures supermodularity in worker-level output: $\partial^2 Z/\partial Y_i \partial Y_j > 0$ for $i \neq j$. As in Kremer (1993), equilibrium features positive assortative matching on the scalar index Y_i.

Now, we can derive the relationship between the population-level skill correlation ($corr_{\text{pop}}$) and the within-firm skill correlation, which we'll denote with $corr_{\text{firm}}$, after workers sort into firms. In Appendix B we solve for the equilibrium, within-firm correlation in skills and obtain:

$$corr_{\text{firm}} = -\exp\left(-\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right) < 0\,. \tag{7}$$

In this example, the within-firm skill correlation is always negative regardless of the population-level correlation. The sharpness of this result stems from the assumptions of (multiplicative) complementarity among all workers within the firm and perfect positive assortative matching. As we highlighted above with Figure 5, this leads to firms where all workers produce on the same isoquant, which, by definition, yields a negative skill correlation. This is a stark example of "Berkson's paradox" (Berkson 1946).

Certainly, the presence of substitutable workers within firms or imperfect positive assortative matching across firms would attenuate this effect.[18] Thus, in practice, within-firm skill correlations need not always be negative.

Still, this model highlights the importance of understanding how workers sort into different levels of aggregation when predicting skill correlations. Under common scenarios, the model indicates that as narrower groups of workers are selected for inclusion in an automation experiment, their skill correlations will become more negative. This, in turn, implies that the choice of scope for an automation experiment can have an important effect on the skill correlation of the workers in the experiment and, in turn, the sign of the inequality effect.

5. Discussion

We have four main results. First, the inequality effects of automation depend on the interaction between skill correlations and the technology's capability relative to worker skills, rather

[18]For example, with random matching, the within-firm and population-level skill correlations will become equal.

than on the workforce's absolute skill level. Second, these effects are often non-monotonic: as the technology improves, inequality may first decrease and then increase, or vice versa. Third, the fastest path to equality (with respect to non-automatable tasks) is balanced technological progress across tasks. Fourth, positive assortative matching within narrowly selected groups can generate negative, within-firm skill correlations even when the population-level correlation is positive.

Practically, this all implies that the same automation technology can reduce inequality in one experimental sample but increase it in another, even if both samples share identical production functions and individual-level skill distributions, so long as their underlying task-level skill correlations differ. Moreover, because of the non-monotonic effects of automation improvements, an experiment showing an inequality reduction today may not imply an inequality reduction tomorrow. Thus, the mix of seemingly inconsistent results documented in automation experiments thus far (Table 1) can fully be rationalized without the introduction of any behavioral biases.[19] A pair of seemingly similar automation experiments will vary in their economic impact because of: (*i*) the task-level skill correlation of workers, (*ii*) which task is (or tasks are) being automated, and (*iii*) the capability of the automation technology relative to the workers' skill levels.

For researchers planning future automation experiments, our framework suggests several priorities. A checklist of considerations motivated by this model include:

- *Document sorting into the experimental population.* Our O-ring analysis highlights the effect that assortative matching can have on skill correlations. Researchers should be clear to report how workers sorted into the population being sampled for the experiment.
- *Measurement of skill correlations.* Skill correlations across workers play a key role in determining inequality effects. While measuring task-specific skills has historically been a challenging exercise, making progress on this front will give empiricists an improved view of the workforce and facilitate better interpretations of the inequality effects obtained.
- *Multiple technology capabilities.* The model highlights how technological progress leads to non-monotonicities. Thus, experimenters could begin to forecast trends by testing for effects using multiple technologies with different capability levels.
- *Task "automatability" measurement.* Since non-automated tasks drive inequality in our model, especially in the long run, documenting which tasks are more or less likely to be automated can improve forecasts beyond the duration of experiments.

Our simple model of automation experiments highlights a unique feature of the workforce: the correlation of workers' skills across tasks. While many discussions of emerging

[19] Of course, that is not to say such biases are irrelevant.

automation technologies focus on how workers with different *levels* of skill are affected, we emphasize here that it is the *correlation* in skills across tasks that is key for understanding how automation affects inequality.

The empirical evidence suggests a negative correlation between physical and mental skills in the population; however, it remains unclear to what extent this is exogenous or an endogenous result of individuals' choices about which skills to acquire and when to acquire them. Digital technologies in the arena of "artificial intelligence" have clearly made a great leap in the automation of mental skills. Thus, for jobs requiring both physical and mental skills, our model suggests initial advances in automation will decrease inequality so long as the task-composition of those jobs remains fixed. Our investigation of SAT math and verbal scores reveals substantial heterogeneity in cognitive skill correlations across industries and occupations; however, all correlations are positive. Therefore, for knowledge workers in the population as a whole, and when technology is not (yet) superhuman, our model predicts automation will disproportionately help lower-skill workers, decreasing inequality. On the other hand, those with jobs like these often find themselves in teams which are highly selected for a particular skill level, for reasons discussed in our O-Ring model extension. In these settings it is possible for the effect of automation on inequality to flip, making extrapolations from the general population to selected teams (or the other direction) problematic. Notably, whatever the skill correlation, there is a clear prediction of what happens as technology becomes increasingly superhuman in a subset of tasks — inequality will increase as the remaining human differences in ability become increasingly magnified.

The model's simplicity is both a strength and a limitation. By abstracting from general-equilibrium adjustments, we isolate the direct productivity channel that experiments measure. This modeling choice makes our predictions directly comparable to experimental results, but it leaves the model silent on longer-run inequality effects in general equilibrium. Incorporating skill correlations into macroeconomic models of automation is a natural next step to explore these broader consequences.

We have also abstracted from heterogeneity in workers' ability to use automation technologies effectively. If "skill in using technology" is itself a task and happens to be correlated with other skills, our model's predictions would become less clear-cut. Recent experimental evidence that some workers perform worse with algorithmic assistance than without it (e.g., Dell'Acqua et al. 2025; Otis et al. 2025) indicates that this is an important extension to consider in future work.

More broadly, our framework suggests that the diversity of automation technologies will play a key role in the evolution of inequality. Technologies that excel at a narrow set of tasks may amplify inequality by primarily benefiting workers who specialize in those tasks or in complementary skills. In contrast, technologies with moderate capabilities across many tasks may compress inequality by helping workers overcome their weakest skills.

This observation motivates further theoretical and empirical work on how the portfolio of available automation tools shapes the returns to various skill combinations.

Finally, our focus on task-level production and skill correlations connects to broader questions about the organization of work. When skills are negatively correlated across workers, firms have an incentive to exploit comparative advantage by assigning each worker to the tasks where their relative skill is highest. When skills are positively correlated, such specialization opportunities are limited, since the same workers tend to outperform others in all tasks. The interaction between skill correlations, automation, and the endogenous organization of production is a promising direction for future research.

References

Acemoglu, Daron, and David Autor. 2011. "Skills, tasks and technologies: Implications for employment and earnings." In *Handbook of Labor Economics*, vol. 4, 1043–1171: Elsevier.

Acemoglu, Daron, and Pascual Restrepo. 2018. "The race between man and machine: Implications of technology for growth, factor shares, and employment." *American economic review* 108 (6): 1488–1542.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and new tasks: How technology displaces and reinstates labor." *Journal of Economic Perspectives* 33 (2): 3–30.

Agarwal, Nikhil, Alex Moehring, Pranav Rajpurkar, and Tobias Salz. 2024. "Combining human expertise with artificial intelligence: Experimental evidence from radiology." *Mimeo*.

Agarwal, Nikhil, Alex Moehring, and Alexander Wolitzky. 2025. "Designing Human-AI Collaboration: A Sufficient-Statistic Approach." *Mimeo*.

Athey, Susan C, Kevin A Bryan, and Joshua S Gans. 2020. "The allocation of decision authority to human and artificial intelligence." In *AEA Papers and Proceedings*, vol. 110: 80–84, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Autor, David H, and Michael J Handel. 2013. "Putting tasks to the test: Human capital, job tasks, and wages." *Journal of labor Economics* 31 (S1): S59–S96.

Autor, David H, Frank Levy, and Richard J Murnane. 2003. "The skill content of recent technological change: An empirical exploration." *The Quarterly Journal of Economics* 118 (4): 1279–1333.

Autor, David, and Neil Thompson. 2025. "Expertise." *Journal of the European Economic Association*: jvaf023.

Baker, Crystal A, Eric Peterson, Steven Pulos, and Rena A Kirkland. 2014. "Eyes and IQ: A meta-analysis of the relationship between intelligence and "Reading the Mind in the Eyes"." *Intelligence* 44: 78–92.

Bárány, Zsofiá L, and Kerstin Holzheu. 2025. "The Two Faces of Worker Specialization." *Mimeo*.

Becker, Gary S, and Kevin M Murphy. 1992. "The division of labor, coordination costs, and knowledge." *The Quarterly Journal of Economics* 107 (4): 1137–1160.

Berkson, Joseph. 1946. "Limitations of the application of fourfold table analysis to hospital data." *Biometrics Bulletin* 2 (3): 47–53.

BLS, (Bureau of Labor Statistics; U.S. Department of Labor). 2024. "National Longitudinal Survey of Youth 1997 Cohort.", Center for Human Resource Research (CHRR), The Ohio State University.

Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond. 2025. "Generative AI at work." *The Quarterly Journal of Economics*: qjae044.

Chen, Zenan, and Jason Chan. 2024. "Large language model in creative work: The role of collaboration modality and user expertise." *Management Science* 70 (12): 9101–9117.

Choi, Jonathan H, Amy B Monahan, and Daniel Schwarcz. 2024. "Lawyering in the age of artificial intelligence." *Minnesota Law Review* 109: 147.

Cui, Zheyuan Kevin, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz. 2025. "The effects of generative AI on high skilled work: Evidence from three field experiments with software developers." *Mimeo*.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan R Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R Lakhani. 2025. "Knowledge Worker Productivity and Quality in Generative AI's Jagged Technological Frontier: Field Experimental

Evidence." *Mimeo*.
Deming, David J. 2017. "The growing importance of social skills in the labor market." *The Quarterly Journal of Economics* 132 (4): 1593–1640.
Dessein, Wouter, and Tano Santos. 2006. "Adaptive organizations." *Journal of Political Economy* 114 (5): 956–995.
Garicano, Luis, and Luis Rayo. 2025. "Training in the Age of AI: A Theory of Apprenticeship Viability." *Mimeo*.
Girsberger, Esther Mirjam, Miriam Koomen, and Matthias Krapf. 2022. "Interpersonal, cognitive, and manual skills: How do they shape employment and wages?" *Labour Economics* 78: 102235.
Guvenen, Fatih, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer. 2020. "Multidimensional skill mismatch." *American Economic Journal: Macroeconomics* 12 (1): 210–244.
Hampf, Franziska, Simon Wiederhold, and Ludger Woessmann. 2017. "Skills, earnings, and employment: Exploring causality in the estimation of returns to skills." *Large-scale Assessments in Education* 5 (1): 12.
Kanazawa, Kyogo, Daiji Kawaguchi, Hitoshi Shigeoka, and Yasutora Watanabe. 2025. "AI, skill, and productivity: The case of taxi drivers." *Management Science*.
Katz, Lawrence F, and David H Autor. 1999. "Changes in the wage structure and earnings inequality." In *Handbook of Labor Economics*, vol. 3, chap. 26, 1463–1555: Elsevier.
Kim, Alex, David S Kim, Maximilian Muhn, Valeri V Nikolaev, and Eric C So. 2025. "AI, Investment Decisions, and Inequality." *Mimeo*.
Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. "Human decisions and machine predictions." *The Quarterly Journal of Economics* 133 (1): 237–293.
Kreitmeir, David, and Paul A. Raschky. 2024. "The Heterogeneous Productivity Effects of Generative AI." *Mimeo*.
Kremer, Michael. 1993. "The O-ring theory of economic development." *The Quarterly Journal of Economics* 108 (3): 551–575.
Lindenlaub, Ilse. 2017. "Sorting multidimensional types: Theory and application." *The Review of Economic Studies* 84 (2): 718–789.
Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional skills, sorting, and human capital accumulation." *American Economic Review* 110 (8): 2328–2376.
Maršál, Ales, and Patryk Perkowski. 2025. "A Task-Based Approach to Generative AI: Evidence from a Field Experiment in Central Banking." *Mimeo*.
Mayer, John D, Richard D Roberts, and Sigal G Barsade. 2008. "Human abilities: Emotional intelligence." *Annu. Rev. Psychol.* 59 (1): 507–536.
Mullainathan, Sendhil, and Ziad Obermeyer. 2022. "Diagnosing physician error: A machine learning approach to low-value health care." *The Quarterly Journal of Economics* 137 (2): 679–727.
Noy, Shakked, and Whitney Zhang. 2023. "Experimental evidence on the productivity effects of generative artificial intelligence." *Science* 381 (6654): 187–192.
Otis, Nicholas, Rowan Clarke, Solene Delecourt, David Holtz, and Rembrand Koning. 2025. "The uneven impact of generative AI on entrepreneurial performance." *Mimeo*.
Roldán-Monés, A. 2025. "When GenAI increases inequality: Evidence from a university debating competition." *Mimeo*.

The Economist. 2025. “How AI will divide the best from the rest.”

Trammell, Philip. 2025. “Workflows and Automation.” *Mimeo*.

Willis, Robert J, and Sherwin Rosen. 1979. “Education and self-selection.” *Journal of Political Economy* 87 (5, Part 2): S7–S36.

Woessmann, Ludger. 2024. “Skills and earnings: A multidimensional perspective on human capital.” *Annual Review of Economics* 17.

Xu, Fasheng, Jing Hou, Wei Chen, and Karen Xie. 2025. “Generative AI and Organizational Structure in the Knowledge Economy.” *Mimeo*.

Zeira, Joseph. 1998. “Workers, machines, and economic growth.” *The Quarterly Journal of Economics* 113 (4): 1091–1117.

Supplementary Appendix

Appendix A. Endogenous Task-level Labor Supply

In the main text, we assume labor is inelastically supplied by all workers at the same cost. Thus, variation in output (and in turn utility) is governed only by the differences in skill endowments. Here, the objective function of the workers in the case of Cobb-Douglas production involves no optimization and is simply given by:

$$\left(\max(A_t\,, \bar{l}_{it}S_{it}) \times \bar{l}_{it-}S_{it-}\right)^{1/2} - 0 \quad \text{where} \quad \bar{l}_{it} = \bar{l}_{it-} = 1\,\forall i\,, \tag{A1}$$

where, as in the main text, we use the notation t to indicate the task being automated and $t-$ to indicate the non-automated task.

Of course, in reality, workers supply effort at some cost and choose how to allocate that effort across tasks in accordance with their expectations. One objective function that more closely captures that reality is:

$$\max_{l_{it}, l_{it-}} \left(\max(A_t\,, l_{it}S_{it}) \times l_{it-}S_{it-}\right)^{1/2} - (l_{it} + l_{it-})^2\,, \tag{A2}$$

which introduces a convex cost of (the sum of) effort and yields closed-form solutions for workers' optimal choices (l^*_{it}, l^*_{it-}) that will depend on their skills and the technology's capabilities.

Solving for workers' optimal choices in Equation (A2) using the same scenarios of skill correlations presented in Table 2 and mirroring the analyses presented in Figure 2 of the main text still yields the result of non-monotonicity in inequality effects as automation improves.

In the case of positive skill correlations, Figure A1A illustrates a pattern very similar to the case of inelastic labor shown in Figure 2A. There is an initial decline in output inequality followed by a long-run increase in output inequality that eventually exceeds the original level of inequality. The transition appears to occur more quickly (i.e., at lower levels of technology capability); however, it is difficult to compare the specific values of the objective functions given that our simple, baseline model has workers facing no costs. Overall, the story remains the same.

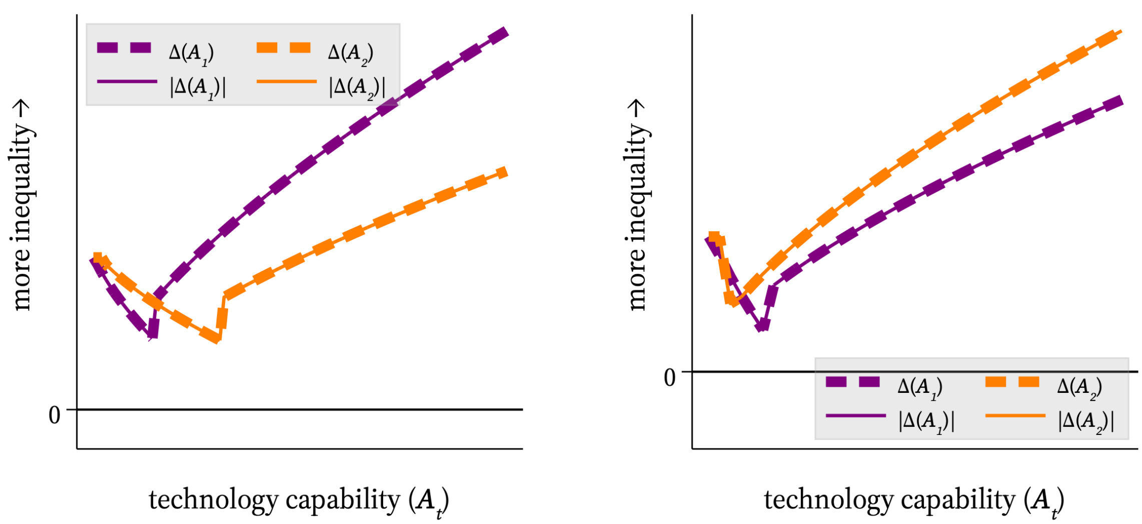
In the case of negative skill correlations, Figure A1B illustrates some patterns that differ somewhat from our baseline scenario shown in Figure 2B. First, the first-difference and absolute measures of output inequality never diverge in this case of endogenous labor supply. This may partly reflect the specific values of the parameters we use here to illustrate the model — the low type never has enough of an absolute advantage in one task to "overtake" the high type. A second difference is that, in this case, the automation of both tasks leads to non-monotonic inequality paths, whereas in the baseline scenario shown in the main text, automation of task 2 only ever increases inequality.

Overall, these analyses reveal that our focal result of inequality effects being non-monotonic in the technology's capability carry through when workers' endogenously allocate their labor across tasks.

FIGURE A1

Inequality and Single-task Automation — Cobb-Douglas Production, Endogenous Labor

(A) Positive skill correlation (B) Negative skill correlation



Note: Plots the level of inequality across the two groups of workers assuming Cobb-Douglas production per first-difference inequality (Δ; solid line) or absolute inequality (|Δ|; dashed line) depending on skill correlations, the task being automated, and the capability of the technology. Based on the case of endogenous labor supply per Equation (A2).

Appendix B. Derivation of Within-Firm Skill Correlations

To summarize the O-ring model setup, each firm employs n workers, and each worker i performs two tasks using their task-specific endowments of skill: $S_{i1}, S_{i2} \in (0,1]$. Worker i's production function is Cobb-Douglas and uses equal shares of both tasks. Following Kremer (1993), firm-level output (Z) is the product of worker-level output: $Z = \prod_{i=1}^{n} Y_i$, which generates positive assortative matching of workers into firms. Skills are distributed such that:

$$\begin{pmatrix} \ln S_{i1} \\ \ln S_{i2} \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu & \sigma^2 & corr_{\text{pop}}\sigma^2 \\ \mu \, , & corr_{\text{pop}}\sigma^2 & \sigma^2 \end{pmatrix},$$

where $corr_{\text{pop}} \in (-1,1)$ is the population correlation of logged skills.

With Cobb-Douglas production, $Y_i = (S_{i1} \times S_{i2})^{1/2}$. Within each firm k, all workers produce output Y_k, so $S_{i1} \times S_{i2} = Y_k^2$ for all workers i in firm k. Taking logarithms of the constraint yields: $\ln S_{i1} + \ln S_{i2} = 2 \ln Y_k$. All within-firm variation in logged skills comes from the difference in logged skill levels, which allows us to write:

$$\begin{aligned} \text{Var}(\ln S_{i1} - \ln S_{i2} \mid Y_k) &= \text{Var}(\ln S_{i1} - \ln S_{i2}) \\ &= 2\sigma^2(1 - corr_{\text{pop}})\,. \end{aligned} \tag{B1}$$

Within each firm, we can express each logged skill levels as: $\ln S_{i1} = \ln Y_k + (\ln S_{i1} - \ln S_{i2})/2$ and $\ln S_{i2} = \ln Y_k - (\ln S_{i1} - \ln S_{i2})/2$. Therefore:

$$\begin{aligned} \text{Var}(\ln S_{i1} \mid Y_k) &= \text{Var}(\ln Y_k + (\ln S_{i1} - \ln S_{i2})/2 \mid Y_k) \\ &= \text{Var}((\ln S_{i1} - \ln S_{i2})/2) \\ &= \frac{\sigma^2(1 - corr_{\text{pop}})}{2}\,, \end{aligned} \tag{B2}$$

and likewise for $\text{Var}(\ln S_{i2} \mid Y_k)$. Now, transforming back to levels, we have $S_{i1} = Y_k \times \exp((\ln S_{i1} - \ln S_{i2})/2)$ and $S_{i2} = Y_k \times \exp(-(\ln S_{i1} - \ln S_{i2})/2)$, where $(\ln S_{i1} - \ln S_{i2})/2 \sim \mathcal{N}(0, \sigma^2(1 - corr_{\text{pop}})/2)$. Therefore:

$$\text{Var}(S_{i1} \mid Y_k) = Y_k^2 \times \exp\left(\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right)\left(\exp\left(\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right) - 1\right), \tag{B3}$$

and likewise for $\text{Var}(S_{i2} \mid Y_k)$.The covariance is then:

$$\text{Cov}(S_{i1}, S_{i2} \mid Y_k) = Y_k^2 \times \left(1 - \exp\left(\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right)\right), \tag{B4}$$

and, finally, the within-firm correlation is:

$$
\begin{aligned}
corr_{\text{firm}} &= \frac{\operatorname{Cov}(S_{i1}, S_{i2} \mid Y_k)}{\sqrt{\operatorname{Var}(S_{i1} \mid Y_k) \times \operatorname{Var}(S_{i2} \mid Y_k)}} \\
&= \frac{1 - \exp\left(\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right)}{\exp\left(\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right)\left(\exp\left(\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right) - 1\right)} \\
&= -\exp\left(-\frac{\sigma^2(1 - corr_{\text{pop}})}{2}\right) .
\end{aligned}
\tag{B5}
$$