Computerization and the Decline of Unincorporated Self-Employment[*]

Yao-Yu Chih[†] Sunhyung Lee[‡] Zexuan Liu[§]

May 8, 2024

Abstract

We investigate the impact of the computer adoption rate, referred to as computerization, on the proportion of unincorporated self-employed (SE) individuals within the US labor markets. The conceptual framework suggests that computerization may either augment or diminish this share. Employing a Bartik instrument approach, we disentangle the causal effect of computerization on unincorporated SE from 1990 to 2010. Our empirical findings indicate that a one percent increase in computerization corresponds to a 0.79 percent reduction in the share of unincorporated SE individuals. Notably, these estimates exhibit variations across industries and metropolitan areas.

Keywords: computerization, entrepreneurs, unincorporated self-employed
JEL Classification: E24, J24, J82, L26, O33.

[*]Acknowledgement: Zexuan Liu acknowledges the financial support of the National Natural Science Foundation of China (grant no. 72103098).

[†]Department of Finance and Economics, McCoy College of Business, Texas State University, San Marcos, TX 78666 USA.

[‡]Corresponding author. Department of Economics, Feliciano School of Business, Montclair State University, 1 Normal Avenue, Montclair, NJ 07043 USA. Email: lees@montclair.edu

[§]Advanced Economic Research Institute, Nanjing Audit University, Nanjing, China. Email: zxliu@nau.edu.cn

1 Introduction

The 1990s and 2000s witnessed an unprecedented surge in technological advancements fueled by the widespread adoption of computers in workplaces, commonly referred to as computerization. Research by Krueger (1993), Autor et al. (1998), Autor et al. (2003), Acemoglu and Autor (2011), Autor (2013), Acemoglu and Restrepo (2018), and others extensively document the profound impact of computerization on workers, wages, and productivity. With the emergence of new technological paradigms such as automation, robotization, and generative artificial intelligence (AI), the need for further exploration into related themes becomes increasingly pertinent and compelling. Despite variations in the experiences associated with each technological wave, a consistent aspect of technological change lies in its capacity to disrupt business dynamics in the US, potentially yielding both beneficial and detrimental outcomes. For instance, automation often replaces routine labor (Autor; 2015; Acemoglu and Restrepo; 2020), yet simultaneously, the least experienced workers in call centers stand to benefit significantly from advancements in generative AI (Brynjolfsson et al.; 2023). This dual effect of technology on low-skilled labor underscores its profound implications for entrepreneurial human capital.

In this paper, we thoroughly examine the impact of computerization on the proportion of unincorporated self-employed (SE) individuals—encompassing both low-skilled laborers and small-scale entrepreneurs—in the United States, spanning from 1990 to 2010. Our study follows an organized approach. Initially, we focus on a distinct demographic group and meticulously analyze data to suggest the diminishing trend in the share of unincorporated SE individuals. Second, we introduce a robust conceptual framework to illuminate how technological advancements can augment or diminish unincorporated SE individuals within the workforce. Third, utilizing comprehensive empirical methods, we demonstrate that the escalation in computerization within our designated timeframe directly correlates with the decline in the prevalence of unincorporated SE. Finally, we examine the broader implications of our findings, particularly within the context of US business dynamism, and contemplate the implications concerning emerging technological paradigms. By adhering to this organized framework, we endeavor to provide a comprehensive understanding of the intricate relationship between computerization and the dynamics of unincorporated SE, thereby offering valuable insights for policymakers, economists, and stakeholders invested in the future trajectory of the US labor market.

Our conceptual framework clarifies the nuanced interplay between computerization and the evolution of unincorporated SE. Technological advancements wield a dual-edged sword, presenting both opportunities and challenges for this demographic. On one hand, the surge in productivity spurred by computerization often displaces workers, precipitating unemployment. Yet, within our conceptual framework, this displacement can catalyze a transition towards self-employment as individuals seek alternative avenues for livelihood, thereby fostering an upsurge in the proportion of entrepreneurs. We term this phenomenon the "restructuring" effect. Conversely, the productivity gains predominantly accrue to large incorporated entities equipped to harness capital more efficiently than their unincorporated counterparts. Consequently, these larger firms may absorb or outcompete smaller businesses, reducing the share of entrepreneurs. We label this as an "efficiency-augmenting" effect. It is a complementary mechanism to the prior research such as Salgado (2020), Kozeniauskas (2022), and Jiang and Sohail (2023), which suggests the concept of skill-biased technological change leading to a decline in the proportion of entrepreneurs. This decline stems from the wage increase for high-skill workers, who synergize with computerization. By delineating these intricate dynamics, our framework enhances the comprehension of how computerization influences the landscape of unincorporated SE.

Following our conceptual exploration, we empirically examine the heterogeneity across workers within commuting zones (CZs) in the US to discern which opposing effects predominate. Given the balanced consideration within our conceptual framework of the restructuring and efficiency-augmenting effects, clarifying the net outcome of these potentially counteracting dynamics becomes imperative. Our methodology and empirical framework closely align with seminal works such as Autor and Dorn (2013), particularly in constructing instruments based on foreign data. Specifically, we instrument US exposure to computerization within CZs using a Bartik-type (or shift-share) measure fashioned from the computerization processes observed in twelve European nations. Notably, Bartik instruments have gained widespread adoption in recent literature, exemplified by works like Goldsmith-Pinkham et al. (2020) and Derenoncourt (2022). Moreover, our study rigorously validates the effectiveness of our Bartik instrument through tests recommended by Borusyak et al. (2022), affirming its robustness and reliability for our analytical pursuits.

The first-stage estimates for the instrumental variable show the substantial predictive power of the European-country instrument for computerization in the US. When introducing the IV estimation, the negative relationship between computerization and the share of unincorporated SE remains robust. The second stage estimates suggest that each one percent increase in computerization precipitated a significant 0.79 percent decrease in the share of unincorporated SE. These estimates affirm that, within our framework, the potency of the efficiency-augmenting effect outweighs that of the restructuring effect. Remarkably, our analysis reveals that this adverse relationship is particularly pronounced in metropolitan areas, where computerization correlates with an additional 0.304 percentage point reduction in the proportion of unincorporated SE. This heightened spatial impact can be attributed to the concentration of skilled labor forces in urban settings, amplifying the dominance of the efficiency-augmenting effect. Our empirical findings describe a compelling narrative: the ascendancy of efficiency-driven mechanisms over restructuring dynamics in shaping the landscape of unincorporated SE within the US, with metropolitan regions as focal points of this phenomenon.

We conduct the following robust analysis to ensure the validity of our main findings and consider possible heterogeneity within our sample. First, we further distinguish by sub-samples to examine how unincorporated SE across industries respond to different degrees of computerization. For sectors with higher computerization levels, such as manufacturing, wholesale trade, and financial sectors, we continue to find a contraction in the share of unincorporated SE because of adopting computers in the workplace. However, the relationship is rather weak for industries with relatively low computerization levels, such as personal services and retail trade. The results confirm our prediction that the marginal efficiency-augmenting effect should be weaker for industries with less computerization. In contrast, the efficiency-augmenting effect dominates the restructuring effect for industries with higher levels and longer histories of computerization. Second, we examine two alternative hypotheses that address industrial compositional change and income effects. Our empirical estimates do not support these two explanations as the driving dynamics of the relationship between computerization and the declining share of unincorporated SE. Third, our empirical findings are robust even when considering several possible confounding factors, such as China shock exposure.

Overall, our findings indicate that new technology diminishes low-quality entrepreneurial endeavors. This positive dynamic effect parallels the observations made in Burtch et al. (2018), where adopting new technology was associated with a decrease in unsuccessful crowdfunding campaigns, notably on platforms like Kickstarter. To delve deeper into the implications of our results within the context of recent technological advancements and the cultivation of entrepreneurial human capital, we dedicate a separate section for thorough exploration and discussion.

We organize the rest of the paper as follows. Section 2 describes the data and measurement. Section 3 presents our conceptual framework. To test our framework, Section 4 outlines the empirical approach. Section 5 reports estimation results, including Section 5.4, which discusses the relative contribution of our research concerning the recent technological change. Section 6 wraps up the paper.

2 Data and Measurements

2.1 Unincorporated SE

Definition While the term "entrepreneur" is widely recognized, achieving a precise definition of entrepreneurs is not without its complexities. Following Glaeser (2009) and Glaeser et al. (2010), we use self-employment as the prominent proxy for entrepreneurial activity. As discussed in Levine and Rubinstein (2013), self-employment can be further described into two main categories: incorporated vs. unincorporated. This classification is based on a legal definition. Incorporated entities provide a distinct legal identity and limited liability, fostering an environment conducive to entrepreneurial endeavors. This structure enables owners to pursue innovative and riskier ventures, supported by investors willing to finance projects with larger scales and longer gestation periods.

Conversely, unincorporated SE individuals often find themselves unable to bear the associated costs of incorporation. They are prevalent in sectors such as personal services (e.g., barbershops, home cleaning services) and retail trades (e.g., auto dealerships, online retail sales), which typically demand lower skill levels and capital investment, predominantly relying on manual skills. However, unincorporated SE individuals can also be found in other industries, such as insurance brokerage. Characteristically, these businesses are owned and managed by individuals or small family groups, focusing on producing a single product or service, catering to localized markets. Despite their entrepreneurial spirit, the total income of unincorporated SE individuals tends to be lower than that of the average wage worker, as noted by Grilo and Thurik (2008), Hunt (2011), Levine and Rubinstein (2013), and Kahn et al. (2017).

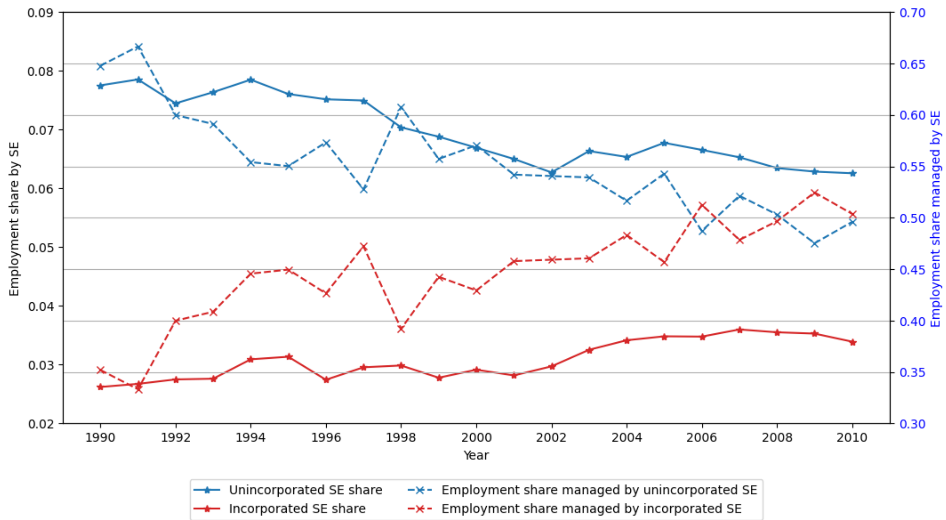
One potential critique of our definition is its susceptibility to variations in state-level legal environments, which impose differing financial and incorporation requirements. To validate the distinctions in our sample between incorporated and unincorporated SE individuals—specifically their smaller scale, lower wealth, and differing sectoral compositions—we conduct a comparative analysis of their distinct characteristics. This analysis is presented in Appendix Table A.1. To gather the necessary data, we turn to the Annual Social and Economic Supplement from the Current Population Survey (CPS) spanning 1990-2010. Our focus is on the civilian non-institutionalized population aged 16-64, actively participating in the labor force and excluding those employed in the agricultural sector. These cross-sectional samples, collected annually in March and appropriately weighted, provide representative insights into this population.

Indeed, the unincorporated SE individuals exhibit lower levels of education, smaller business sizes, and lower income than their incorporated counterparts. Specifically, 45% (28.4%) of the unincorporated (incorporated) individuals have a high school degree as their highest level of education. Nearly 77% (60%) of all unincorporated (incorporated) SE individuals operate businesses with fewer than ten employees. On average, unincorporated (incorporated) SE individuals earn approximately \$38,000 (\$70,000) annually. Moreover, the unincorporated SE individuals are predominantly represented in the personal services and construction sectors, while the incorporated SE individuals are more prevalent in retail/wholesale trade and manufacturing industries.

Decline of Self-Employed? It is crucial to differentiate between entrepreneurs based on their incorporation status because failing to do so could lead to a skewed perception of entrepreneurship in the US. The literature documents that there has been a noticeable decline in the proportion of entrepreneurs in the US since the 1980s. Specifically, self-employed (Kozeniauskas; 2022), self-employed business owners (Salgado; 2020), fast-growing firms (Decker et al.; 2016b), and new businesses (Decker et al.; 2014) have all fallen in the economy, and Jiang and Sohail (2023) even document that the decrease is more pronounced for college graduates. The overall decline in entrepreneurship has an important economic implication: the US economy has become less dynamic. Decker et al. (2016a), Decker et al. (2016b), Decker et al. (2017), and Decker et al. (2020), Salgado (2020), Jiang and Sohail (2023), and many others have raised concern that the decline in dynamism could reduce productivity, job creation, business formation, and even economic growth. As for the causes of the decline, a myriad of literature offers different mechanisms. Kozeniauskas (2022) points to increasing entry costs, primarily from regulation and technological changes. Salgado (2020) attributes to the skill-biased technological change and the decrease in the price of capital. Similarly, Krusell et al. (2000) focus on the complementarity and substitutability of skilled and unskilled labor (respectively) with capital and the decline in the price of capital. Jiang and Sohail (2023) further address the association between the decline in the price of capital goods and the skill premium, using college graduates as a proxy for high-skill labor.[1]

[1]Other reasons include how the aging of the population affects the formation of new businesses (Engbom et al.; 2019) or how the decrease in the labor force growth affects the construction of new companies in the context of a firm dynamics model (Karahan et al.; 2019; Hopenhayn et al.; 2022).

Figure 1: Evolution of Unincorporated vs. Incorporated SE Share, 1990-2010



Notes: This figure plots the annual rate of US unincorporated and incorporated SE shares from 1990 to 2010 using the CPS data.

Evolution of Unincorporated and SE However, this trend might not accurately portray the entire entrepreneurial landscape, as it overlooks the crucial distinction between incorporated and unincorporated SE. Figure 1 illustrates the evolution of US unincorporated and incorporated SE shares from 1990 to 2010. It is important to clarify that the data presented in this figure is sourced from the Current Population Survey (CPS), a widely used monthly survey in the US that offers detailed labor force data, demographic information, insights into program participation, and supplemental topics. Since the 1988 survey, the CPS has differentiated between unincorporated and incorporated SE, reporting them separately. For this figure, we utilize the proportion of workers identified as unincorporated or incorporated SE to illustrate the patterns of occupational choice.

As substantiated by statistical evidence in Appendix Table A.1, incorporated SE tends to have larger firm sizes. Hence, using the number of heads is not an appropriate metric if the goal is to demonstrate how the employment share of firms managed by unincorporated or incorporated SE has evolved at a different rate. Unfortunately, to our knowledge, the employment share of firms concerning whether they are unincorporated or incorporated is not explicitly available in any dataset. Similar to Kozeniauskas (2022), we estimate an employment share to address this limitation. We assume that firms are managed either by incorporated or unincorporated entities. Utilizing data from the CPS, we determine firm size based on the self-employed individual's response to the "*FIRMSIZE*" question, which elicits information about the number of employees categorized into groups such as "under 25 employees," "25 to 99 employees," and so on. We calculate the median value for each category to estimate the average firm size for both unincorporated and incorporated SE individuals. This approach enables us to approximate the employment shares attributed to unincorporated and incorporated SE.

In this vein, Figure 1 presents both the number of heads and the estimated employment shares for unincorporated and incorporated SE. The two solid lines represent the headcounts of the unincorporated and incorporated SE shares. The two dashed lines indicate the estimated employment shares associated with unincorporated and incorporated SE. Since we have assumed that firms are either managed by incorporated or unincorporated entities, the sum of these two dashed lines equals one each year. Within our sample periods, there was a discernible uptick in the proportion of incorporated SE, indicating that the observed decline in overall SE share primarily stems from the decrease in the unincorporated segment. Furthermore, the number of employees overseen by unincorporated SE individuals has dwindled, contrasting with the ascent in employees managed by their incorporated counterparts. In conclusion, our preliminary analysis illustrates a notable decline in the share of unincorporated SE individuals from 1990 to 2010.

Then, what has been driving the persistent decline in the share of unincorporated SE individuals over the past two decades? Anecdotal evidence suggests that small family businesses in the US have been grappling with the challenges posed by the emergence of big-box retail chains like Walmart and online giants such as Amazon.[2] While these narratives may hold some truth at the establishment level, empirical evidence may not entirely support claims that factors such as the presence of big-box retail chains like Walmart and online giants like Amazon definitively led to the decline of small businesses in the US since 1990. Notably, research by Sobel and Dean (2008) suggests that Walmart stores have little long-term impact on small businesses' overall number and profitability, based on a sample spanning all 50 states from 1985 to 2002.

At the same time, even though changes in consumption behavior, offshoring, subsidies, and tax breaks may have contributed to the decline of small businesses in the US, Kuratko and Audretsch (2022) argue that the proliferation of disruptive technology is the primary factor shaping the evolution of small entrepreneurs in the US. Among various technological advancements, computerization stands out as a transformative force that has dramatically reshaped the business landscape. Therefore, we aim to investigate whether computerization can account for the continuous decline in the share of unincorporated SE individuals.

2.2 Computerization

Definition We construct a measure of computerization at the local economy level, i.e., commuting zone, between 1990 and 2010. In the spirit of Basso et al. (2020), we combine information on computerization at the industry level as measured by 1989 (matched with the 1990 census IPUMS) and 2001 (matched with the 2000 census IPUMS) Computer and Internet Use supplement survey. The Computer and Internet Use supplement survey has been included in the CPS since 1984. Specifically, we use the variable "Directly use a computer at work" to measure computerization. The "Directly use a computer at work" item was addressed to civilians aged 15 or older to determine whether the respondent uses a computer at work. As suggested by the survey, we use the "Computer and Internet Use supplement weight" to aggregate the individual-level data. Then, we impute this measure at the local level, exploiting the industrial structure of each commuting zone in 1980 (Tolbert and

[2] Crow, David and James Fontanella-Khan. 2020. "Is this the end for America's mom-and-pop stores?" *Financial Times*, August 27. https://www.ft.com/content/92427a94-ee5e-486c-9f6b-9e11e8362f41

Sizer; 1996).

Our definition of computerization may face criticism due to the ambiguity surrounding computer usage at work. For example, a receptionist at McDonald's may use a computer daily, but this does not necessarily imply that the restaurant relies heavily on computer technology. Conversely, Amazon's delivery driver may not use a computer daily despite the company's business model being primarily based on online logistics. Due to data limitations, the computer and internet use supplement survey targets individuals rather than firms. However, our analysis focuses on variations at the CZ level rather than at the individual or firm level. Therefore, as long as the survey is representative of the firm, the CZ-aggregated computerization rate should accurately reflect the computer adoption rate. In other words, while Amazon's delivery driver may not use computers daily, there will likely be a sufficient number of programmers and office employees who have been surveyed and provided positive responses to the question regarding computer usage.[3]

2.3 Commuting Zone Analysis

Background Large sample sizes are essential for analyzing changes in labor market composition at the detailed geographic level. Our analysis draws on the census Integrated Public Use Micro Samples (IPUMS) for 1990 and 2000 and the American Community Survey (ACS) for 2010. The 1990 and 2000 census samples are 5 percent of the US population, and the ACS sample is 1 percent of the population. Tolbert and Killian (1987) first introduced the concept of commuting zones (CZs) as a proxy measure for local labor markets because neither counties nor metropolitan statistical areas (MSAs) are geographically defined based on economic activities. Thus, neither a county nor an MSA is the most appropriate representation of a local labor market. Since the CZs identified by Tolbert and Killian (1987) are not fully compatible with the definitions in the 1990 census, Tolbert and Sizer (1996) modified this approach using county-level commuting data from the 1990 census and created 741 clusters of counties characterized by strong commuting ties within CZs and weak commuting ties across CZs. Autor and Dorn (2013) argue that one of the advantages of using CZs over counties is that CZs are primarily based on economic geography rather than incidental factors such as the minimum population. Our analysis includes the 722 CZs that cover the mainland of the United States (both metropolitan and rural areas).

Our sample consists of individuals between the ages of 16 and 64 working full-time the entire year preceding the survey. Residents in institutional group quarters, such as prisons and psychiatric institutions, are dropped in the sample, along with unpaid family workers.[4] Labor supply is measured by the number of weeks worked multiplied by the usual weekly hours. All calculations are weighted by the census sampling weight multiplied by the labor supply weight and a weight derived from the geographic matching process described in Autor and Dorn (2013).[5]

[3]In Appendix A.3, we provide further analysis to ensure that this assumption holds within our conceptual framework.

[4]We exclude unpaid family workers from the sample, following the methodology of Autor and Dorn (2013). In Appendix A.2, we offer additional justification for this decision.

[5]Following Autor and Dorn (2013), this weight is used to aggregate the individual-level survey data to the CZ-level data. The census sampling weight is defined in the census IPUMS. Because the basic units in the IPUMS are public use micro areas (PUMAs) (1990, 2000, and 2010), county groups (1980), and state economic areas (SEAs) (1950), which are

Variable Construction It is crucial to highlight that our primary data source for estimations relies on the decennial Census survey, which forms the basis of our analysis regarding regional variations across CZs over the decades. Although the CPS offers yearly data, its reporting unit is too large to consistently measure unincorporated SE at the CZ level between 1990 and 2010. Additionally, the annual CPS data has limitations concerning the availability of the computerization measure derived from the Computer and Internet Use Supplement Survey, which was only accessible for specific years within our sample period (1989, 1993, 1997, 2001, and 2003).

Due to the data constraints of the Computer and Internet Use Supplement Survey, obtaining specific CZ or county data for respondents was not feasible, making it challenging to determine the computerization rate directly for each CZ. Consequently, we employed an imputation method to estimate the CZ-level computerization rates using available local industrial shares and national computerization data.

Since the survey provides each respondent's residential state, we calculated the state-level computerization rate. Our approach aggregated the imputed CZ-level computerization rates to the state level by comparing these rates with the computerization rates derived from the raw survey data. This methodological approach aimed to capture significant variations in computerization across different CZs, as reflected in discrepancies between the imputed and raw computerization rates at the state level.

In Appendix Figure A.2, we compare the state-level imputed computerization rates plotted against the survey-based rates. This visualization demonstrates a positive correlation between our imputed rates and the raw computerization rates at the state level. Furthermore, data points do not predominantly cluster on either side of the 45-degree line, indicating that our imputed figures do not systematically skew towards over- or underestimating the computerization rate from the survey-based measure.

It's important to note that while our method may underestimate computer usage in CZs with intensive computer use and overestimate it in less intensive CZs, the data indicate no systematic bias towards either type of CZ. This finding strongly supports the unbiased nature of the average effect presented in our manuscript, which remains the primary focus of our study.

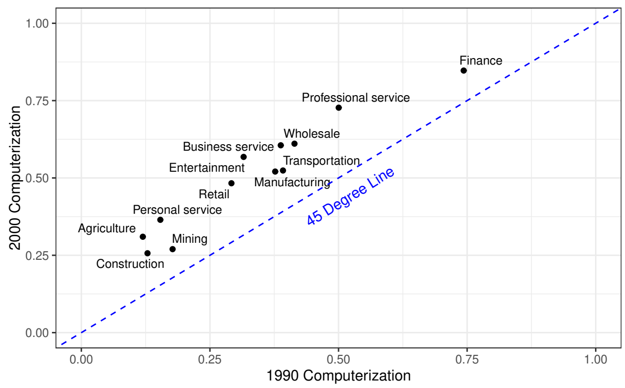
2.4 The Rise of Computerization and the Decline of Unincorporated SE

Our survey-based data encompasses a total of 12 industries. Analyzing this dataset, we observe a notable rise and variability in computerization, which is a significant source of identification. Figure 2 visually depicts the statistics for computerization and unincorporated SE by decade.

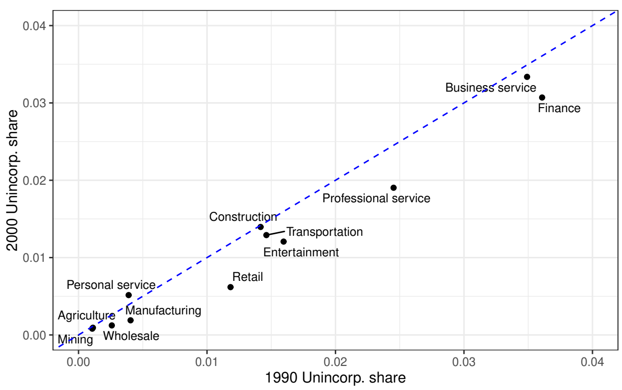
Rise of Computerization Figure 2a examines computerization across various industries. Notably, sectors such as finance, professional services, and wholesale trade exhibit the highest levels of computerization, while industries like agriculture, mining, and construction show comparatively lower

combinations of counties, a CZ may be comprised of multiple PUMAs, county groups or SEAs. A PUMA, county group, or SEA can also be mapped to different CZs. The weight derived from this geographic matching process can be loosely understood as the percentage of a geographic unit that falls into a specific CZ.

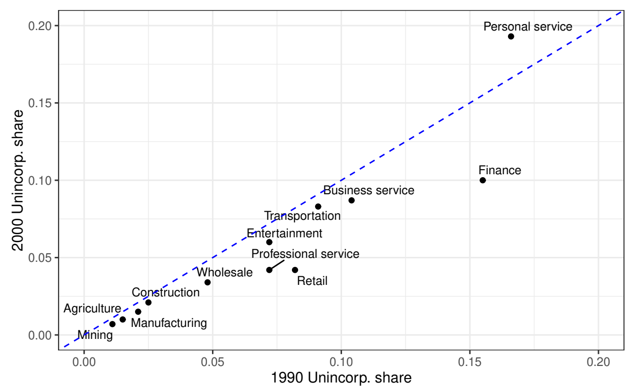
Figure 2: The Rise of Computerization and the Decline of Unincorporated SE



(a) Computerization across Industries



(b) Unincorp. SE Share in National Employment



(c) Unincorp. SE Share by Industrial Employment

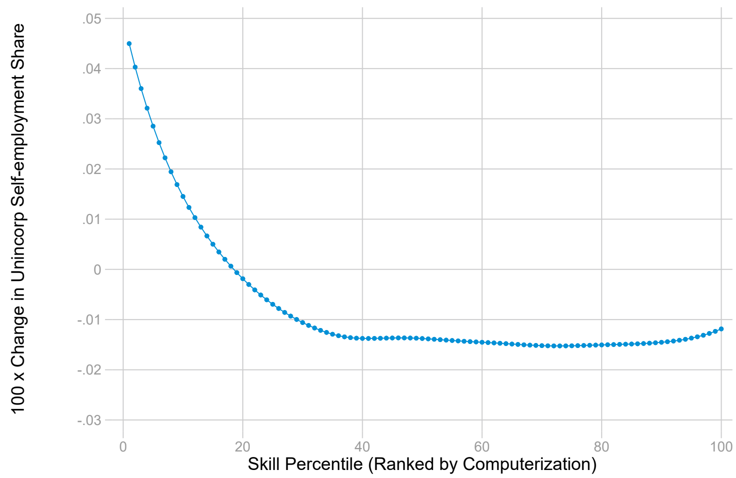
Notes: This figure plots the different rates of computerization and the share of unincorporated SE in national employment and industrial employment by 1990 and 2000 for the twelve industries observed in our survey-based data.

levels. This initial analysis suggests that certain industries are more exposed to computerization than others.

When comparing computerization trends across decades, we observe a consistent upward trajectory. All 12 industries analyzed lie above the 45-degree line, indicating an overall increase in computerization over the decades across all sectors. Particularly noteworthy are the significant percentage point increases in computerization observed in industries such as entertainment services (from 31% to 57%), professional services (from 50% to 73%), business services (from 39% to 61%), and personal services (from 15% to 36%) from 1990 to 2000. This analysis highlights that service-related sectors have experienced the most substantial changes among the 12 industries examined. In total, our survey-based data suggests an increase in overall computerization.

Decline of Unincorporated SE Figure 2b illustrates the share of unincorporated SE among nationally employed laborers in each industry. Meanwhile, Figure 2c showcases the share of unincorporated SE in each industry among nationally employed laborers within the same industry, highlighting the contribution of unincorporated SE to aggregate employment at the industrial level. Both

Figure 3: Unincorporated SE Share by Computerization Percentile



Notes: This figure plots the change in the unincorporated SE share by the percentile of the imputed computerization.

figures demonstrate a decrease in the share of unincorporated SE over the decade for most industries, as evidenced by their positions below the 45-degree lines. The only exception is the personal services industry, which makes intuitive sense as it is relatively more difficult for computerization to replace the personal services industry than other industries.

3 Conceptual Framework

In the preceding section, we noticed an association between the increase in computerization and the decrease in unincorporated SE during the 1990s to 2000s through survey-based data analysis. To examine deeper into this relationship, we will now describe how the share of unincorporated SE changes across various levels of computerization, as depicted in Figure 3. We rank all industries defined in the IPUMS according to their computerization levels in 1990 and construct the smoothed difference in the share of unincorporated SE between 1990 and 2010 for each computerization percentile. The findings reveal a strikingly heterogeneous growth rate in the share of unincorporated SE concerning the level of computerization. Industries in the lowest quantile exhibit a significant increase in the percentage of unincorporated SE individuals. However, this growth pattern does not extend to the upper quantiles, where the share declines. Serving as descriptive evidence, this figure underscores the heterogeneous association between computerization and the change in the share of unincorporated SE individuals. This heterogeneous impact implies that a single effect does not drive the relationship but arises from the dynamic interplay between at least two-dimensional effects.

Given the heterogeneous relationship observed between computerization and the share of unincorporated SE, we present a conceptual framework to formulate hypotheses regarding the role of computerization and its counteracting forces on the share of unincorporated SE.[6]

While existing literature has identified various factors contributing to the decline in entrepreneurship, such as the slowdown in labor force growth, heightened friction or start-up costs impeding firm adaptation, and the ongoing aging of the population,[7] We posit that computerization may be a significant factor driving this decline, with a particular focus on unincorporated SE. Therefore, our research aligns closely with studies investigating the relationship between skill-biased technological change and the decline of entrepreneurship.

The fundamental conceptual framework adopted from this research asserts that computerization allows computers to replace routine-intensive labor while concurrently boosting the productivity of higher-skilled workers.[8] For instance, computerization can facilitate automation on production lines, initially operated by routine-intensive workers, but subsequently requiring additional roles for computer engineers, data scientists, and programmers to oversee and maintain the automated processes. Prior studies, such as those by Krusell et al. (2000), Autor and Duggan (2003), Autor and Dorn (2013), and Acemoglu and Autor (2011), have provided theoretical insights and empirically demonstrated evidence supporting this trend. These studies primarily concentrate on general capital improvements that drive skill-biased technical change and emphasize the employment polarization within the US labor market.[9]

More recent papers have extended the skill-biased technical change framework into studying the decline of entrepreneurship. Notably, Jiang and Sohail (2023) present findings indicating the skill-biased nature of the decline in entrepreneurship. Their research highlights that the decrease in entrepreneurship has been notably more pronounced among individuals with higher skill levels. Utilizing a model of occupational choice with work heterogeneity based on Lucas Jr (1978), they demonstrate that an increase in the skill premium, resulting from skill-biased technological advancements, has a minimal impact on lowering the overall entrepreneurship rate. Instead, the decline in entrepreneurship is predominantly attributed to skill-neutral technological changes and an increasing supply of college graduates. Particularly, the rising skill premium diminishes the average productivity of entrepreneurs as the composition of entrepreneurs shifts away from skilled individuals towards those with lesser skills. This shift occurs because the earnings of skilled individuals grow at a faster rate than those of entrepreneurs, thus discouraging the pursuit of entrepreneurship among the skilled population.

[6]In Appendix B.1, we provide more discussions of our conceptual framework's relation to the skill-biased technological change literature.

[7]For further exploration of the slowdown in labor force growth, readers are directed to Hopenhayn et al. (2022) and Karahan et al. (2019); for insights into heightened frictions or start-up costs hindering firm adaptation, refer to Decker et al. (2020); and for discussions on the ongoing aging of the population, see Bornstein et al. (2018) and Engbom et al. (2019).

[8]Acemoglu and Restrepo (2018) argue that recent advancements in AI, in conjunction with advances in big data and machine learning, have enabled computerization to automate complex tasks traditionally performed by high-skilled workers, potentially displacing both high and low-skilled workers. However, our study's sample period spans from 1990 to 2000, during which computer automation primarily focused on routine tasks. Thus, high-skilled workers were largely insulated from automation. Therefore, the narrative remains applicable.

[9]Recent research (Lefter et al.; 2011; Mishel et al.; 2013; Hunt and Nunn; 2022) has raised empirical concerns regarding the concept of employment polarization within the framework of skill-biased technical change. However, the endogenous skill supply setting continues to provide a more comprehensive and realistic environment compared to the conventional canonical model, and the growing demand for high-skilled individuals remains evident.

In the same vein, Salgado (2020) contends that the decrease in entrepreneurship can be traced back to two technological factors that have heightened the rewards for highly skilled labor: skill-biased technical change and a decrease in the cost of capital. Through the utilization of a quantitative model of entrepreneurial decision-making,[10] Salgado (2020) demonstrates the significance of these factors in elucidating the rise in the skill premium, which in turn can suggest the decrease in the share of entrepreneurs. More precisely, Salgado (2020) reveals that skill-biased technical changes explain fifty percent of the reduction in the entrepreneur share, with the remaining fifty percent equally distributed between the decrease in capital goods costs and the increase in the availability of high-skilled labor.

Kozeniauskas (2022) presents a dynamic occupational choice model[11] specifically crafted to encapsulate theories explaining the decline in entrepreneurship. Among four potential explanations,[12] Kozeniauskas (2022) identifies that the most compelling reason for the decrease in the share of individuals who are entrepreneurs is the rising entry and fixed costs attributed to both increasing regulations and changes in technology. Conversely, skill-biased technical change reallocates entrepreneurship towards less educated individuals but doesn't seem significant in explaining the overall shift in entrepreneurship at an aggregate level. Although productivity gains by non-enterprise firms contribute to the decline in entrepreneurship, they are not the primary force compared to the rising entry and fixed costs.

Our empirical hypothesis closely aligns with the pioneering studies mentioned earlier on the decline of entrepreneurship. Instead of constructing a calibration model grounded in the extensively explored and documented framework of skill-biased technical change, as investigated by various versions of the occupational choice model, we adopt the conclusions drawn from these pioneering studies as hypotheses for empirical testing. Our focus is on empirical investigation, particularly on establishing a causal relationship between computerization and the decline in the share of unincorporated SE using the shift-share IV approach. For our analysis, we utilize the census Integrated Public Use Micro Sample (IPUMS) and the American Community Survey (ACS) to construct commuting zones as a proxy measure for local labor markets, as opposed to household-level data from Panel Study of Income Dynamics (PSID) or Current Population Survey (CPS) commonly used in the literature. Therefore, our study serves as a complementary contribution to further support prior research.

[10] The model presented by Salgado (2020) expands upon the span-of-control model introduced by Lucas Jr (1978) by incorporating two distinct types of workers: high and low skills. This extension includes elements such as time-varying entrepreneurial ability, asset accumulation, and idiosyncratic labor risk, which are commonly considered in entrepreneurial decision-making frameworks like those proposed by Cagetti and De Nardi (2006) and Quadrini (2000).

[11] Within the dynamic general equilibrium model of occupational choice, as proposed by Kozeniauskas (2022), agents are assumed to possess the capability to engage in either low or high-skill work, along with entrepreneurial productivity. In each period, these agents decide whether to be inactive in the labor force, work as employees, or operate a business as entrepreneurs. Additionally, a non-entrepreneurial sector exists. All businesses utilize the same production technology, characterized by fixed and entry costs, and incorporate capital along with two types of labor as inputs. Skill-biased technical change is represented through changes in capital prices, while the superstar firm hypothesis is depicted by an increase in the relative productivity of non-entrepreneurial firms, accompanied by fixed and entry costs.

[12] According to Kozeniauskas (2022), four possible reasons for the decline in entrepreneurship include: (1) skill-biased technical change causing changes in wage, (2) increases in regulatory burdens increasing fixed and entry costs for businesses, (3) the adoption of technology shifting costs towards fixed and entry expenses, and (4) change in technology enhancing the relative productivity of larger firms compared to smaller ones.

In addition, distinguishing our research from the aforementioned pioneer studies on the decline of entrepreneurship, we specifically concentrate on unincorporated SE, which represents a subset of entrepreneurship. Examining unincorporated SE through empirical investigation is crucial due to the distinct differences it holds in comparison to incorporated SE individuals, despite both falling under the umbrella of entrepreneurship. Illustrated in Figure 1, the noteworthy decline in overall entrepreneurship over the past two decades coexists with a contrasting upward trajectory in incorporated SE. Furthermore, a significant discrepancy emerges in terms of employment structures, as unincorporated SE demonstrates a lower likelihood of having paid employees. To provide context, in 2014, approximately 41% of incorporated SE individuals had at least one paid employee, whereas only 13% of unincorporated SE individuals exhibited this characteristic, based on data from the CPS.[13] Given the larger scale of employment, incorporated SEs may share certain characteristics more closely aligned with non-entrepreneurial firms than unincorporated SEs do. For instance, it may benefit from higher productivity resulting from skill-biased technical change. Therefore, concentrating exclusively on the analysis of unincorporated SE not only sets our research sample apart from pioneer studies[14] but also helps in understanding the factors behind the decline of SE, free from the influence of information pertaining to incorporated SE.

Drawing from the literature, we argue that computerization may exert two opposing forces — namely, a restructuring effect and an efficiency-augmenting effect — on the evolution of the share of unincorporated SE.

In terms of the restructuring effect, as outlined in the employment polarization theory, routine-intensive workers and computerization serve as substitutes. With the advancement of computer technology, tasks previously carried out by routine-intensive employees become automated (Krusell et al.; 2000; Autor and Duggan; 2003; Autor and Dorn; 2013; Acemoglu and Autor; 2011; Acemoglu and Restrepo; 2018), leading to a decreased demand for such workers and, consequently, a form of restructuring unemployment. We hypothesize that some displaced workers may choose to pursue SE outside traditional non-entrepreneurial corporate structures. Thus, in occupations with a higher concentration of routine-intensive workers, we anticipate a rise in the prevalence of unincorporated SE.

Furthermore, both Salgado (2020) and Kozeniauskas (2022) assert that the decrease in computer capital price offers advantages to self-employed individuals by streamlining operations and boost-

[13] Based on data from the 2014 CPS, although unincorporated SE constitutes 63% of the total reported self-employment share, its contribution to overall hiring is limited, accounting for only 24%. Among self-employed individuals who engage in hiring, the average number of paid employees differs significantly between incorporated and unincorporated SE, with figures standing at 10.2 and 5.6, respectively, in the year 2014.

[14] Salgado (2020) constructs four nested classifications of entrepreneurship as follows: (1) all households who are business owners (referred to as "business owners"), (2) business owners who actively worked for their business during the year ("active business owners"), (3) households who are both business owners and actively worked for their business, and whose head is self-employed ("self-employed business owners"), and (4) the subset of self-employed business owners who hold a managerial or professional occupation (referred to as "entrepreneurs"). Kozeniauskas (2022) defines an entrepreneur as someone self-employed, with a business comprising at least ten employees, aged between 25 and 65, and not employed in the agriculture or government sectors. Although our study focuses only on unincorporated SE, it does not strictly adhere to Kozeniauskas (2022)' criteria as we do not impose restrictions regarding the number of employees, age groups, or sectors within our sample. With just 13% of unincorporated SE individuals hiring paid employees and an average of 5.6 employees per business, it implies that a significant portion of the sample studied in this research differs from that of Kozeniauskas (2022) and Salgado (2020).

ing efficiency across various tasks. These tasks may encompass data analysis, financial tracking, client or customer communication, and marketing endeavors. Leveraging computerization, self-employed individuals can automate repetitive tasks, access valuable online resources, collaborate with remote teams, and utilize a range of software tools to optimize their business processes. As a result, entrepreneurs can allocate more time and resources toward innovation, strategic planning, and business expansion, thereby increasing the likelihood of profitability and sustainability in their ventures. This effect is particularly pronounced in occupations with a higher concentration of routine-intensive workers, as the rise in wages for routine-type workers is less likely to surpass the increase in profits achievable through self-employment due to the restructuring within an economy. As a result, the likelihood of transitioning from marginal self-employment to becoming a worker is reduced.

We encapsulate the restructuring effect and its positive impact of computerization on the share of unincorporated SE in Hypothesis 1.

Hypothesis 1 The restructuring effect from computerization increases the unincorporated SE share. First, firms invest in computerization, displacing tasks previously handled by routine-intensive employees. Some individuals who were displaced choose to adapt their skills and enter into unincorporated SE. Second, the decreased price of computer capital benefits unincorporated SE individuals, as they can leverage computers to enhance their operations, thereby boosting the prospects of profitability and sustaining their businesses. Both points suggest that growth in the share of unincorporated SE is expected in occupations with a higher concentration of routine-intensive workers.

In terms of the efficiency-augmenting effect, Salgado (2020) and Kozeniauskas (2022) contend that technological advancements enhance the relative productivity of larger firms,[15] leading to an increased demand for high-skilled labor and, subsequently, higher wages. This trend may discourage individuals, especially those with high skills, from opting for self-employment. Given that a significant portion of unincorporated SE businesses are small-scale enterprises unable to compete with larger firms in terms of wage offerings for highly skilled workers, we hypothesize that high-skilled self-employed individuals may opt to close their businesses and seek employment with larger firms. Therefore, in occupations with a higher concentration of abstract-intensive workers, we anticipate a decrease in the share of unincorporated SE.

Furthermore, Kozeniauskas (2022) contends that the change in technology is one of the factors driving the increase in entry and fixed costs within production.[16][17] The rise disproportionately affects smaller businesses, including unincorporated SE individuals. Higher entry and fixed costs decrease the payoff from being unincorporated SE, resulting in fewer individuals pursuing such occupations and consequently reducing the share of unincorporated SE businesses.

In summary, non-entrepreneurial firms benefit more from the efficiency-augmenting effect due

[15] As per Autor et al. (2020), larger companies possess a strategic edge in harnessing new technologies, thanks to their scale and superior financial access. Furthermore, advancements in technology empower consumers to make more informed comparisons regarding prices and quality, thereby favoring the most efficient firms.

[16] Aghion et al. (2022); De Ridder (2024); Hsieh and Rossi-Hansberg (2023) argue that the growing utilization of IT technology has raised the entry and fixed costs of firms.

[17] Another factor is the increasing regulation, as noted by Kozeniauskas (2022).

to their comparatively higher productivity growth and lower burden of entry and fixed costs compared to unincorporated SE. We encapsulate the efficiency-augmenting effect and its negative impact of computerization on the share of unincorporated SE in Hypothesis 2.

Hypothesis 2 The efficiency-augmenting effect from computerization decreases the unincorporated SE share. First, computerization enhances the productivity of non-entrepreneurial firms, thereby heightening the demand for highly skilled labor engaged in abstract-intensive tasks, leading to increased wages. This prompts more individuals engaged in unincorporated SE to exit and opt for positions within non-entrepreneurial firms. Second, the growing reliance on computerization has increased the entry and fixed component of firm costs. Consequently, fewer individuals, particularly those with high-skilled workers, opt for self-employment. Both factors indicate that in occupations with a higher concentration of abstract-intensive workers, a decline in the share of unincorporated SE is expected.

The net effect of computerization on changes in the share of unincorporated SE is influenced by both restructuring and efficiency-augmenting factors. Recent research indicates a decrease in entrepreneurship, suggesting that the efficiency-augmenting effect prevails. However, studies also provide evidence of a shift in entrepreneurship towards individuals with lower levels of educational attainment, indicating the presence of a restructuring effect. In Appendix B.2, we present a simple task-based model developed by Acemoglu and Autor (2011) to illustrate how the mechanisms of the two effects interact and influence the changes in the share of unincorporated SE.

4 Empirical Approach

In this section, we explain the identification strategy to estimate the impact of computerization on the change of the unincorporated SE share at the CZ level.

4.1 Baseline Empirical Strategy

In equation terms, our CZ-level computerization measure is constructed as

$$Computerization_{c,t_0}^{US} = \sum_j \omega_{j,c,t_0} \times \frac{Computerization_{j,t_0}^{US}}{empl_{j,t_0}^{US}} \quad \text{and} \quad \omega_{j,c,t_0} = \frac{empl_{j,c,t_0}}{\sum_j empl_{j,c,t_0}}, \tag{1}$$

where $Computerization_{c,t_0}^{US}$ is the computer adoption rate in a commuting zone c at time t_0, based on the local CZ c's industrial structure (ω_{j,c,t_0}) and the national level of computerization of industry j at time t_0. More specifically, we exploit the information contained in the 1989 and 2001 Current Population Survey (CPS) microdata, as available on IPUMS (Ruggles et al.; 2020) to construct the nationwide industry-level share of computerization at the time of the 1990 and 2000 censuses ($\frac{Computerization_{j,t_0}^{US}}{empl_{j,t_0}^{US}}$). That is, the data on computerization and the share of the unincorporated SE are not collected in the same year but rather have a one-year difference.[18] However, this measurement error is not likely to significantly bias our estimation as computerization is unlikely to change quickly. Specifically, we measure the following regression form:

$$\Delta \text{Unincorporated SE}_{c,t} = \alpha + \beta \times \text{Computerization}^{US}_{c,t_0} + \eta_s + \tau_{t_0} + \tau_{s,t_0} + \varepsilon_{c,t_0}, \tag{2}$$

where the dependent variable is the change in the unincorporated SE share over the decade t_0 to t_1 in a commuting zone c,[19] Computerization$^{US}_{c,t_0}$ is the commuting zone's average computerization at the start of the decade; η_s represents state fixed effects, measuring any time-invariant differences across states; and ε_{c,t_0} is the unobserved error terms. We also include time dummies (τ_{t_0}) and, as a robustness check, state-specific linear time trends (τ_{s,t_0}) to capture any dynamic effects of computerization on the share of unincorporated SE. Standard errors are clustered at the commuting zone level.[20]

The coefficient β in Equation (2) captures the marginal effect of a unit increase in computerization on a commuting zone's unincorporated SE share growth. The most straightforward strategy for estimating β is the ordinary least squares (OLS) approach. However, in our setup, a potential problem of the fixed-effects approach is that it cannot address time-varying unobserved heterogeneity related to the unincorporated SE share.

In addition, the growth of unincorporated SE may affect computerization. For instance, incorporated companies can outsource some of their businesses to unincorporated SE if the cost is lower than that of performing tasks within the firm, even after computerization. This limits the firm's production boundary and lowers computerization. This reverse causality also causes a downward bias of the coefficient β in the fixed-effects model.

Moreover, the OLS estimates may have a downward bias because of potential measurement errors in the computerization indicator. The computerization measure depends on the local commuting zone's industrial composition and the national-level industrial computer adoption rate. One caveat concerning this approach is that it ignores CZ-level characteristics other than industrial composition associated with the likelihood of adopting computers. For instance, commuting zones with higher labor costs are likelier to computerize codeable tasks than areas with relatively low labor costs, *ceteris paribus*. This predicted computerization measure may underestimate the true value in metropolitan areas while overestimating computerization in non-metropolitan areas.

There is also a potential concern regarding the computerization variable in Equation (2). The ω_{j,c,t_0} term represents the share of employment in a specific industry within a CZ at time t_0. If the decline in the share of entrepreneurs varies across sectors, it could artificially create a negative

[18]Because the CPS supplemental survey is not carried out yearly and the computer usage information is derived from the 1989 and 2001 CPS supplemental microdata, we do not have computerization that perfectly matches the 1990 and 2000 censuses.

[19]The Δ*Unincorporated SE*$_{c,t}$ is calculated as $\frac{\text{Unincorp SE}_{c,t_1} - \text{Unincorp SE}_{c,t_0}}{\text{Unincorp SE}_{c,t_0}}$, and the *Unincorporated SE*$_{c,t}$ is defined as the share of unincorporated SE over the employed labor.

[20]In addition, the state-specific time trends control for whether there is any pretreatment deviation in the outcome variable that is correlated with the treatment; for example, if states that have a high computerization rate are also more likely to increase their unincorporated SE share, then this confounding variation can be appropriately controlled for by adding state-specific time trends (Page et al.; 2005; Addison et al.; 2009; Allegretto et al.; 2011; Meer and West; 2016).

correlation between computerization and the share of entrepreneurs due to its effect on ω_{j,c,t_0}. Additionally, if the decline in entrepreneurs leads to a shift from less "computerizable" sectors to more "computerizable" ones, the latter's share would increase over time, further generating a negative correlation. This can potentially bias the results through the same directional change of the share of entrepreneurs and the ω_{j,c,t_0} because the share of entrepreneurs is part of the CZ's local industrial structure.

Therefore, we address these issues using an instrumental variable approach in the next section that exploits historical variation in the local industry structure.

4.2 Alternative Specification: Instrumental Variable Estimates

To identify the component of computerization driven by technological innovation, we instrument the US exposure to computerization using an analogous measure constructed from computerization in twelve European countries.[21] The computerization data for European countries are obtained from the European Working Conditions Survey (EWCS), which includes over 12,500 employees and self-employed labor.[22]

Specifically, following Autor et al. (2013) and Acemoglu and Restrepo (2020), we construct a shift-share type of instrument, Computerization$^{EU}_{c,t_0}$,[23] as follows:

$$Computerization^{EU}_{c,t_0} = \sum_j \sum_k \omega_{c,j,k,1950} \times \frac{Computerization^{EU}_{j,k,t_0}}{empl^{EU}_{j,k,t_0}} \quad \text{and} \quad \omega_{c,j,1950} = \frac{empl_{c,j,k,1950}}{\sum_j \sum_k empl_{c,j,k,1950}}, \tag{3}$$

where the subscript k denotes occupation, $\omega_{c,j,k,1950}$ represents the local industry-occupation group share. Appendix Table C.1 and Table C.2 provide the mapping we manually created and used for harmonizing the 1990 and 2000 EWCS industry and occupational codes to the 1950 US Census codes. The industry and occupational classifications are only available in one digit of the EWCS data. In 1990, there were 12 occupational groups and 9 industrial groups, while by 2000, those numbers shifted to 10 occupational groups and 11 industrial groups. Therefore, $\omega_{c,j,k,1950}$ includes 64 and 78 unique industry-occupation group exposure shares. We provide further details of our matching procedure in Appendix C.1.

Equation (3) suggests that the non-US exposure to computerization differs from Equation (1) in two respects. First, in place of US industrial computerization (Computerization$^{US}_{j,t_0}$), it uses the level of computer penetration in other high-income European countries (Computerization$^{EU}_{j,k,t_0}$) by

[21] The twelve other European countries are Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, and United Kingdom.

[22] Specifically, each member state had approximately 1,500 employees, except Luxembourg, which had 500 employees. We provide further details on the EWCS data in Appendix C.1.

[23] The survey questions on which we rely on to generate the computerization measure are "Does your main paid job involve - working with computers: PCs, network, mainframe?" (1990 EWCS) and "What is the main activity of the company or organization where you work ?" (2000 EWCS). Both are categorical questions, and we define computerization as at least 25% or more time of an individual's main paid job involves working with computers. In the robustness test, we upgrade the threshold to 100% of the time, and the results are consistent.

industry-occupation groups. Second, this expression uses the local employment level in 1950 instead of the current period's employment share. We use a four-decade-lagged employment level because the current local industry structure and the unobserved labor demand shock within a CZ are most likely correlated. Thus, the use of lagged employment will mitigate the simultaneity bias. It is also worth noting that even when considering different sub-periods of our baseline model, we keep the employment share constant at the year 1950 to avoid the mechanical correlation between the local industry structure and labor demand shocks.

The shift-share instrument we discussed in Equation (3) is introduced by Bartik (1991) and then popularized in Blanchard and Katz (1992). In this setup, Goldsmith-Pinkham et al. (2020) argue that the instrument is equivalent to using the exposure shares (lagged regional industry-occupational composition) as instruments. Thus, the exogeneity condition depends on exposure shares without imposing any explicit assumption of the shock variable (industry-specific penetration of computers in European countries). In contrast, Borusyak et al. (2022) show that under certain assumptions, the identification of this instrument is achieved when the shocks are as good as random, while exposure shares are allowed to be endogenous. Therefore, the validity of our instrument depends on some exogeneity conditions about the regional industry-occupational composition in 1950, computerization in European countries, or both.

There are several threats to our identification approach because the validity of our instrument depends on whether the industry-occupation-specific computer adoption in twelve European countries is orthogonal to the unobserved error term in Equation (1). For instance, if the average productivity growth of high-skill workers in twelve European countries is more than that for the US, some high-tech service jobs will be moved out of the US. That would lead to a fall in wage growth of high-skill workers in the US, leading to a fall in demand for low-skill in-person service jobs. Although we cannot rule this possibility out, evidence from these studies suggests that the productivity growth of high-skill workers from 1990 to 2010 in the US is much more rapid than in most of the major European countries.

Another possible threat to the instrument validity is that the rise in Chinese labor productivity may be associated with the ongoing decline in the cost of computer production in both the US and European countries through trade with China. Hence, Chinese export growth may adversely affect employment in some labor-intensive industries. To address this concern, in a robustness exercise, we include the change in import exposure per CZ as a proxy for the local labor market exposure to import competition and find that the results are robust, which supports the validity of the exclusion restriction. Furthermore, we allow state-specific time trends to test whether there exist any dynamic effects of computerization on the change of unincorporated SE share. The results from this robustness exercise are also quite reassuring.

Table 1 presents the first-stage estimates for the instrumental variable estimation strategy. The coefficient in column (1) of Panel A is the estimated $\text{Computerization}_{c,t_0}^{US}$ for which the local exposure is at the 1-digit industry (9 and 11 unique groups) level. In column (2), the local exposure share is constructed based on our proposed industry-occupation groups. After allowing for the computerization across industry-occupation groups, the coefficients drop from 1.05 to 0.87. In column (3),

Table 1: 2SLS IV First Stage Regression Results and Correlation between Predicted and Observed Computerization

Panel A: First stage estimation of $\text{Computerization}^{US}_{c,t_0}$					
	(1)	(2)	(3)	(4)	(5)
	All sample			1990	2000
$\text{Computerization}^{EU}_{c,t_0}$ (industry-occupation group)		0.870***	0.637***	0.529***	0.645***
		(0.066)	(0.067)	(0.070)	(0.086)
$\text{Computerization}^{EU}_{c,t_0}$ (industry level)	1.050***				
	(0.165)				
Unemployment rate			0.216	0.444***	0.177
			(0.139)	(0.144)	(0.192)
Percentage Female employment/Pop			0.210***	0.252***	0.089
			(0.047)	(0.051)	(0.066)
Age 65+/Pop			-0.325***	-0.262***	-0.375***
			(0.060)	(0.056)	(0.066)
Share of minimum wage workers			-0.044	-0.277***	-0.411***
			(0.045)	(0.070)	(0.131)
R^2	0.923	0.948	0.958	0.795	0.766
Observations	1,444	1,444	1,444	722	722

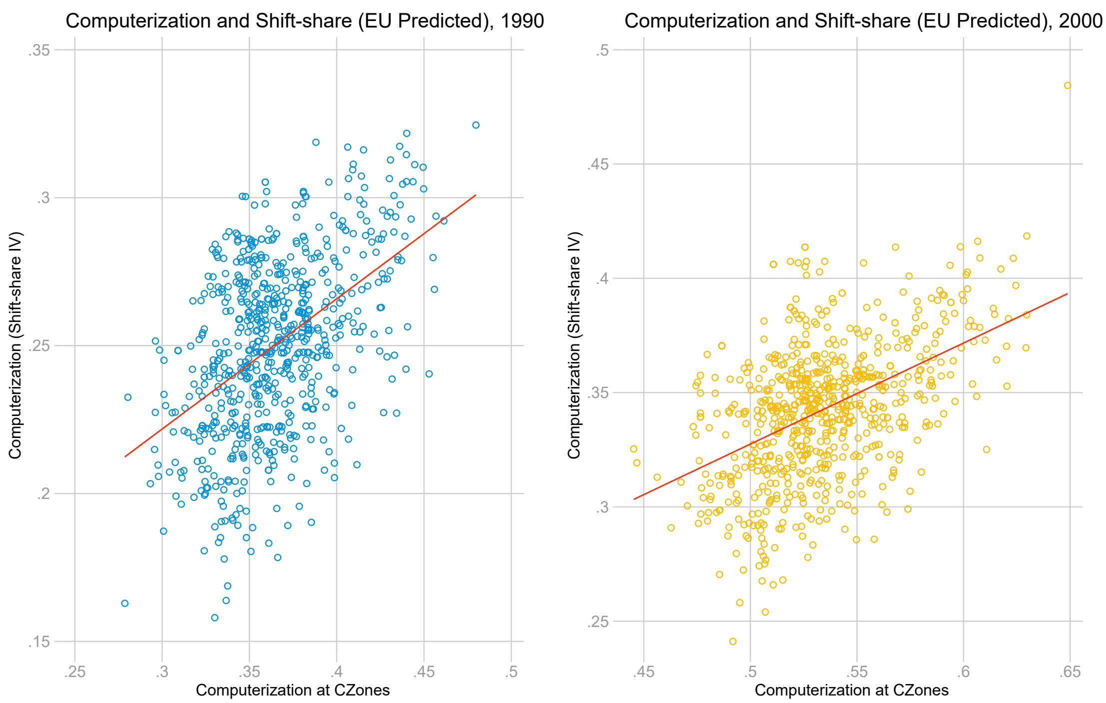
Panel B: Correlation between $\text{Computerization}^{EU}_{c,t_0}$ *and* $\text{Computerization}^{US}_{c,t_0}$				
	$\text{Computerization}^{EU}_{c,t_0}$ (industry-occupation group)		$\text{Computerization}^{EU}_{c,t_0}$ (industry level)	
	(1)	(2)	(3)	(4)
	1990	2000	1990	2000
$\text{Computerization}^{US}_{c,t_0}$	0.721	0.740	0.482	0.617

Notes: This table reports the 2SLS IV first stage regression results (Panel A) and the correlation statistics between the predicted and observed computerization variables (Panel B). All the estimations models control for state and time fixed effects and are weighted by the share of each commuting zone in national population in each sample year, and the standard error is clustered at the commuting zone level. * Denotes significance at a 10% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

we show that the conditional correlation between $\text{Computerization}^{US}_{c,t_0}$ and $\text{Computerization}^{EU}_{c,t_0}$ in our sample period is 0.64 and statistically significant at the 1 percent level. This strong association reveals the substantial predictive power of the European-country instrument for the penetration of computers in the US. In columns (4) and (5), we find similar positive and solid correlations for the sub-periods of 1990 and 2000. We graphically demonstrate these results in Figure 4. The solid lines in Figure 4 for 1990 and 2000 denote the linear fit of the scatter-plot and indicate that the association between $\text{Computerization}^{US}_{c,t_0}$ and $\text{Computerization}^{EU}_{c,t_0}$ is more substantial in 2000 than 1990.

Figure 5 shows the industry-level computer adoption correlation between the US and the twelve European countries. The blue circles and gray triangles denote specific industries in 1990 and 2000, respectively. There were nine industries in 1990 and twelve in 2000. Consistent with our presumption that technological improvements drive US industry trends in computerization, there is a positive correlation between the adoption of computers in the European countries and the US (also see Table 1 Panel B). Figure 5 also reveals a significant heterogeneity of computer use across in-

Figure 4: 2SLS First Stage Regression: Predicted Computerization and Observed Computerization in 1990 and 2000

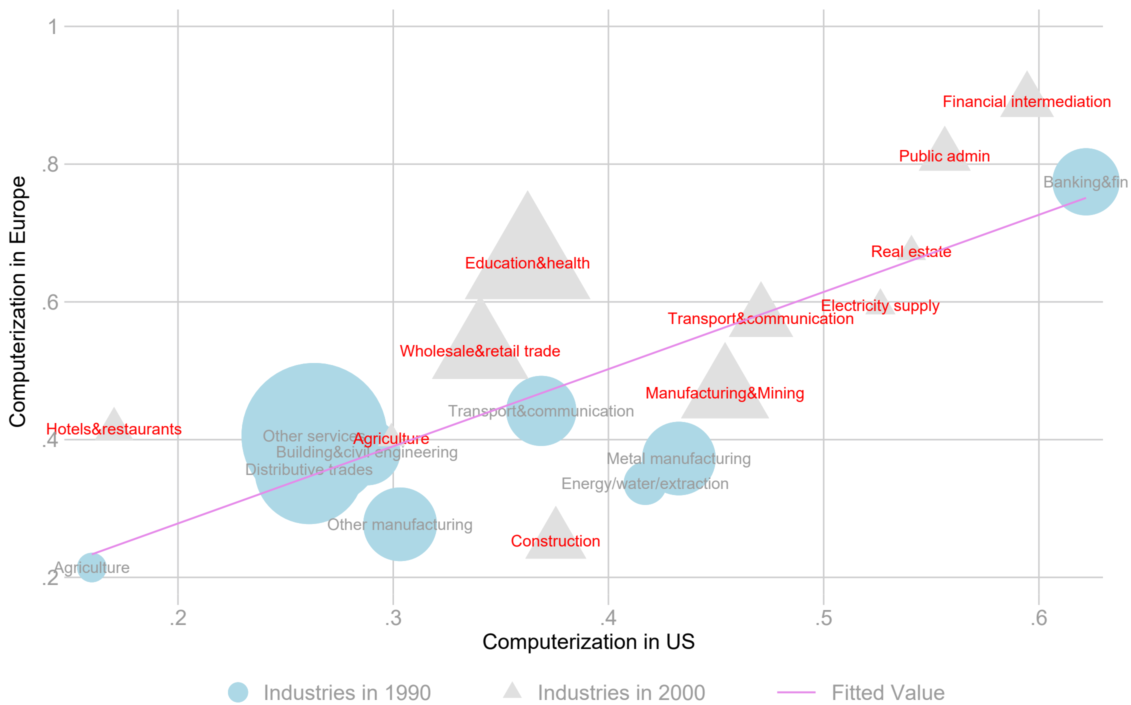


Notes: This figure depicts the 2SLS IV first stage regression results for 1990 (left panel) and 2000 (right panel), respectively. Each data point denotes community zones. The shift-share (EU predicted Computerization) for 1990 and 2000 are obtained from Table 1. The solid lines in both panels represent the best linear fit of the data.

dustries. While some industries such as financial institutions, banks, public administration, real estate, transportation, and communication have more than 40% of computer usage, others such as hotels and restaurants, construction, agriculture, distributive trades, and other manufacturing experienced modest computer adoption in both the United States and Europe.

While we have provided an *a priori* argument supporting the instrument validity in shift-share IV, it is essential to conduct several diagnostics to examine the extent of shock variation and assess this assumption's falsifiability. These post hoc assessments will help us evaluate the *ex-post* plausibility of the instrument. Both Goldsmith-Pinkham et al. (2020) and Borusyak et al. (2022) present an econometric framework for conducting valid shift-share IV inference and testing. Our shift-share IV bears resemblance to the approach adopted by Autor et al. (2013) and Acemoglu and Restrepo (2020). In this setting, the shifts (shocks) are tailored to address a specific research question, i.e., the growth of computerization. In contrast, the shares are considered generic, as they have the potential to capture an observation's exposure to multiple shocks. Goldsmith-Pinkham et al. (2020) found that this form of exposure shares (lagged employment shares) is deemed implausible as instruments based on various balance and over-identification tests. Hence, we will utilize the identification approach proposed by Borusyak et al. (2022). This approach hinges on the quasi-random assignment

Figure 5: Association Between Industry Specific Computerization in European Countries and the US for 1990 and 2000



Notes: This figure plots the association between an industry-specific computerization rate between the US and EU for 1990 (circles) and 2000 (triangles).

of shifts (shocks), allowing the exposure shares to be endogenous. Furthermore, we will assess the validity of the identifying assumption proposed by Lamadon et al. (2022).

Borusyak et al. (2022) provide a novel econometric framework for shift-share instrumental variable regressions in which there are tests for determining whether the "shift (shock)" is exogenous. Specifically, the shift would have to be from the quasi-random assignment. The "share" component can be endogenous within their framework. Even this is so; we use the 1950s share as part of our IV, which is unlikely to be affected by computerization, meaning that our share IV is most likely exogenous. To determine the exogeneity of the shift part of our IV, we follow the series of tests suggested in Borusyak et al. (2022). First, Table 2 reports summary statistics for the computerization shock with importance weights. Column (1) shows that the distribution of the computerization shock has a mean value of 0.325 and a standard deviation of 0.252 with an interquartile range of 0.385. The inverse Herfindahl index (HHI) is relatively regular, with 48.4 across industry-by-occupation-by-period cells and 11.6 when aggregated at the broadest level.[24] The most significant shock weights are 7.2% across industry-by-occupation-by-period cells and 16.5% across the broader level. These

[24]The inverse of Herfindahl index (HHI), computed as $1/\sum_{n,t} s_{nt}^2$, where s_{nt} is the industry-occupation level weights, serves as an indicator of industry concentration, and aligns with the concept of effective sample size.

values translate to a reasonably sizable degree of variation at the industry level, satisfying the baseline assumptions of the quasi-experimental framework. As suggested in Borusyak et al. (2022), we report the inverse of the HHI of shock-level average exposure to demonstrate an adequate sample size. The subsequent F-statistic values reported in our IV estimation will provide a formal statistical test of the power of this shock variation. Column (2) summarizes the distribution of within-period computerization shocks. Borusyak et al. (2022) states that the distribution would be leveraged by an assumption of conditional quasi-experimental assignment. We regress the computerization shocks on the period fixed effects with s_{nt} weights and obtain the summary statistics values of these residuals. The standard deviation of these shock residuals has a value of 0.248, and the interquartile range is 0.368, which are only slightly smaller than those from column (1), which are 0.252 and 0.385, respectively. This confirms that there is sizable residual shock variation, even conditional on the period. This is another evidence that we satisfy the baseline assumptions of the quasi-experimental framework regarding the effective sample size.

Table 2: Summary Statistics for the Industry-Occupational Level Shocks

	(1)	(2)
Mean	0.325	0
Standard deviation	0.252	0.248
Interquartile range	0.385	0.368
Specification		
All industry-occupations	✓	✓
Residualizing on period FE		✓
Effective sample size ($1/HHI$ of s_{nt} weights)		
Across industry-occupations and periods	48.388	48.388
Across industry (level1) groups	11.579	11.579
Largest s_{nt} weight		
Across industry-occupations and periods	0.072	0.072
Across industry (level1) groups	0.165	0.165
Observation counts		
No. of industry-occupations shocks	142	142
No. of industry (level1) groups	11	11

Notes: This table summarizes the distribution of computerization shocks g_{nt} across industry-occupation group n and periods t. Shocks are measured as the share of computerization in European countries. All statistics are weighted by the average industry-occupation group exposure s_{nt}. Both columns include all the industry-occupation groups. We follow Borusyak et al. (2022) to report the effective sample size (the inverse renormalized Herfindahl index of the s_{nt} weights at the industry-occupation-period level and at the level of industrial (level 1) groups, along with the largest s_{nt}.

Next, we need to test whether the shocks are sufficiently mutually uncorrelated. Hence, we assess the correlation patterns of shocks across industries using the industry classifications and periods of the pooled cross-section available within our data. We calculate our shocks' intra-class correlation coefficients (ICCs) within different industry groups. This would be equivalent to a standard parameter clustering method. Using a random effects model, we have the following decomposition of residual within-period shock variation:

$$g_{nt} = \mu_t + \alpha_{industry} + \delta_n + e_{nt}, \tag{4}$$

where g_{nt} is the industry-occupation level shocks, $n = 1, \dots, N$, μ_t is the period fixed effects, $\alpha_{industry}$ is the time-varying random effect computed by our industry group specification, and δ_n is a time-invariant industry-occupation random effect, which spans across 71 industry-occupation pairs. We obtain the estimates using a maximum likelihood approach from a hierarchical linear model. Table 3 reports the estimated ICCs using Equation (4). Again, it shows the summary of the share of the overall shock residual variance because of each random effect. Our survey data for computerization asks whether a workplace has a computer adoption rate of 25% or a rate of 100%. Column (1) reports the estimate from the computerization at 25%. The ICC for the larger industry group is 0.051 with no statistical significance. This means there is less evidence for clustering shocks at the higher industry classification. The ICC for the industry-occupation pair is 0.84, which is statistically significant at the 1% level. This means there is evidence for moderate clustering of shock residuals at this classification level. Column (2) reports the estimate from the computerization at 100%. A similar pattern is found for this level of computerization as the findings from column (1). Our ICCs analysis supports the assumption that shocks are mean-independent across the industry-occupation pair clusters. It indicates that our sample would be sufficient to cluster standard errors at that classification level, and we would still have an adequate, effective sample size.

Table 3: Shock Intra-class Correlations

	(1) Computerization at 25%	(2) Computerization at 100%
Shock ICCs		
industry (level1) groups	0.051	0.045
	(0.033)	(0.038)
industry-occupations	0.840***	0.837***
	(0.032)	(0.035)
Period means		
1990s	0.351***	0.091***
	(0.038)	(0.014)
2000s	0.418***	0.127***
	(0.037)	(0.016)
No. of industry-occupation periods	142	142

Notes: This table reports intra-class correlation coefficients following Borusyak et al. (2022), estimated from the hierarchical model. Estimates come from a maximum likelihood procedure with an exchangeable covariance structure for each industry-occupation group random effect and with period fixed effects. Robust standard errors are reported in parentheses.

As a final step, we conduct a falsification test of the computerization shock orthogonality. The main focus of our study is the regional balance tests and pre-trends analysis. As suggested in Borusyak et al. (2022), we implement the falsification tests in a two-step process. First, potential proxies for the unobserved residual are regressed on the instrument. One caveat is that we use exposure-robust inference that accounts for the inherent dependencies of our data. We implement our regression at the shock level to obtain the exposure-robust standard errors. Second, we regress each potential confounder on the shift-share instrument and the share-weighted average of period effects. The shift-share instrument is normalized to have a unit variance. The share-weighted average of period effects is equivalent to the period-interacted sum of shares because our data has incomplete shares.

Table 4 reports the estimation results of our regional balance tests. Within our commuting zone population, we have the control variables for the share of foreign-born, unemployed, employment among women, people over the age of 65, workers who receive the minimum wage, and the average offshorability index of occupations. These variables are not statistically significant, implying that we find no meaningful associations between these control variables and the shift-share instrument within periods. Moreover, we conduct a regional pre-trends analysis as suggested in Lamadon et al. (2022). We follow the same specification as the regional balance test in the previous rows for this test. The main difference is that we regress the pre-trend variables (the share of unincorporated SE in the 1970s and the 1980s) on the shift-share instrument. Our finding from the pre-trends analysis shows that both coefficients for the 1970s and the 1980s are statistically insignificant. It implies no meaningful association between the shift-share instrument and the share of unincorporated SE in both decades. In all, we fail to reject imbalance in all of the eight potential confounders at conventional levels of statistical significance. Based on the findings presented in Tables 2–4, we find supporting evidence for interpreting the shift-share IV in this study as leveraging quasi-random variation in industry-occupation-specific computerization shocks.

5 Main Results

5.1 Computerization and the unincorporated SE Share

5.1.1 Detailed OLS Estimates

We begin by estimating Equation (2) to provide baseline estimates for the effect of computerization on the unincorporated SE share in Table 5. These estimates mainly provide a general starting point and reference for the results of computerization in different industries in the following subsection.

Column (1) shows the estimated coefficients of computerization for the two periods (i.e., 1990s and 2000s) pooled OLS model, with state and year fixed effects. Standard errors are clustered at the commuting zone level to account for any systematic correlations across commuting zones. By construction, the mean of computerization is 0.480 in the sample period, and the population-weighted 80/20 percentile range is 18.5 percentage points ($Computerization^{P20}$ = 0.402 and $Computerization^{P80}$ = 0.586). The estimates in column (1) suggest that a commuting zone at the 80th percentile of com-

Table 4: Shock Balance Tests

Balance variable	Coef.	SE
Panel A: Regional balance		
Start-of-period % of foreign-born population	0.026	(0.066)
Start-of-period % of unemployed	-0.004	(0.007)
Start-of-period % of employment among women	0.052	(0.034)
Start-of-period % of age 65+	-0.029	(0.024)
Start-of-period % of minimum wage workers	-0.043	(0.045)
Start-of-period average offshorability index of occupations	0.205	(0.156)
Panel B: Pre-trends Analysis		
Δ share of unincorporated SE, 1970	-0.104	(0.260)
Δ share of unincorporated SE, 1980	0.100	(0.106)

Notes: Panel A reports coefficients from regressions of commuting zone-level covariates on the shift-share instrument, controlling for period indicators. The regional balance variables vary across the two periods. Panel B reports coefficients from pre-trend regressions on the shift-share instrument, controlling for period indicators. The industry-occupation-clustered exposure-robust standard errors are reported in parentheses and obtained from equivalent industry-level IV regressions as described in Borusyak et al. (2022). Independent variables in both panels are normalized to have a variance of one in the sample.

puterization experienced a 10.915-percentage-point (18.5×-0.590) more significant contraction of the unincorporated SE share between 1990 and 2010 than did a 20th-percentile commuting zone.

In columns (2), (3), and (4), we include the beginning-period routine-intensive and abstract-intensive occupational shares, which are the occupational structure measures at the CZ level proposed by Autor and Dorn (2013). We discuss the construction of routine- and abstract-intensive occupation share and more detailed implications in Appendix D.1. Computerization will replace routine-intensive workers, shifting wage workers to unincorporated SE. In column (2), the coefficient of the routine-intensive occupation share (RSH) is positive and significant, which supports the restructuring effect hypothesis (Hypothesis 1). Moreover, including RSH in column (2) does not change the significance level of computerization.

Furthermore, we include the abstract-intensive occupational share (ASH) in column (3), and the coefficient is negative and significant. Specifically, computerization is expected to augment the productivity/efficiency of wage workers, especially those conducting more abstract-intensive type tasks. This efficiency-augmenting effect should attract more workers seeking to benefit from economies of scale into incorporated firms, thereby decreasing the share of unincorporated SE, thus lending support to Hypothesis 2. Note that including ASH in the model changes the significance level of computerization, suggesting that the efficiency-augmenting effect fully mediates the declining unincorporated SE. Column (4) includes the beginning periods of both RSH and ASH, and the results are consistent with those in columns (2) and (3).

One concern is that the downfall of unincorporated SE is driven by the legal or tax treatment that may shift self-employed labor from unincorporated to incorporated. Failing to account for this

Table 5: Computerization and the Change in the Unincorporated SE Share within CZs

	$10 \times$ *annual percentage change in share of unincorporated SE*									
	OLS							2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Computerization	-0.590***	-0.806***	0.179	0.149	-0.586***	-0.568***	-0.546***	-0.791***	-0.737***	-1.090***
	(0.111)	(0.126)	(0.240)	(0.234)	(0.111)	(0.113)	(0.109)	(0.155)	(0.152)	(0.334)
Share of routine-intensive occupations		0.642***		1.008***						
		(0.197)		(0.203)						
Share of abstract-intensive occupations			-0.576***	-0.806***						
			(0.145)	(0.155)						
Δ Incorp SE share					-0.086***					
					(0.023)					
Δ Share of labor hired by Incorp SE					0.018					
					(0.020)					
(Δ import from China to US)/ worker						0.005**			0.004	
						(0.002)			(0.004)	
Computerization $\times$ Metropolitan										-0.304**
										(0.118)
Metropolitan										0.183***
										(0.060)
Constant	0.247***	0.187***	0.087	-0.071	0.252***	0.217***	0.175**	0.357***	0.315***	0.529***
	(0.066)	(0.069)	(0.083)	(0.091)	(0.066)	(0.069)	(0.080)	(0.087)	(0.089)	(0.178)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State specific time trend							Yes			
Underidentification test										
Kleibergen-Paap statistic								81.594	19.507	40.123
Weak identification test										
F statistic								174.983	46.646	38.246
critical 10% value								16.38	7.03	7.03
Observations	1,444	1,444	1,444	1,444	1,444	1,444	1,444	1,444	1,444	1,444
R^2	0.169	0.176	0.181	0.197	0.180	0.173	0.387	0.167	0.171	0.172

Notes: All the estimations models control for state and time-fixed effects and are weighted by the share of each commuting zone in the national population in each sample year and the standard errors clustered at the commuting zone level. Column (7) additionally includes the state-specific time trend for the robustness. The definition of Δ import from China to the US)/ worker follows Autor et al. (2013). The Metropolitan dummy is defined as whether a metropolitan city exists within the commuting zone. * Denotes significance at a 10% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

shift leads to an overestimation of the impact of computerization. As shown in Figure 1, the share of incorporated SE quickly rises in some sub-periods (e.g., the 2000s); however, the change is modest overall. In addition, the period-by-period examination does not provide descriptive evidence: the incorporated SE was reasonably stable in the 1990s, but the unincorporated SE share dropped significantly. Furthermore, the incorporated SE share rose relatively fast in the 2000s, while the unincorporated SE share did not appear to plunge simultaneously.

We test whether the increase of incorporated SE share leads to the contraction of unincorporated SE in column (5). The rising incorporated SE share is negatively correlated with the change in the unincorporated SE share. Notably, the coefficient of computerization is mainly unaffected by controlling for changes in the incorporated SE share,[25] which does not justify the "shift" between

[25] This is mainly because computerization is not significantly correlated with changes in incorporated SE. Therefore, any tax treatment or regulatory policies that may increase the share of incorporated SE is not significantly associated with computerization. The result is robust to the use of the beginning-of-period incorporated SE share. The estimation results are available upon request.

registered corporations and the unincorporated SE.

Next, the initial computerization level is associated with an industrial mix or occupational structure in a commuting zone. The industrial or occupational structure is subject to shocks correlated with both the level of computerization and the missing unincorporated SE. Following Autor et al. (2013), we add the change in import exposure per CZ worker[26] in column (6) as a proxy for the local labor market exposure to import competition, which is commonly known as the "China shock." The import competition shocks US manufacturing employment, which reinforces the restructuring effect. Thus, it causes a significant amount of employed manufacturing labor to be relocated to unincorporated SE. As expected, the import competition positively correlates with the unincorporated SE share at the 5 percent significance level. However, the coefficient of computerization remains unaffected. For the sensitivity test, column (7) includes the state-specific time trend to account for heterogeneous time trends across states and the state and national time fixed effects. The coefficients remain robust and are statistically comparable to those in column (1).

5.1.2 Instrumental Variables Estimates

We report the results based on the instrumental variable estimates in columns (8)-(10) in Table 5. One crucial factor we do not observe is the unincorporated entrepreneur's ability. Although the skill requirements for unincorporated SE are significantly lower than those for incorporated entrepreneurship, computerization can still accelerate the learning process of unincorporated SE individuals to help them better organize their establishments. Failing to consider the learning effect leads to an upward bias of the OLS regression.

Column (8) presents the IV estimates comparable with the OLS estimates in column (1). The IV coefficient for computerization is -0.791, which is a more significant negative impact than -0.590 of the OLS coefficient. This initial comparison shows that the OLS specification underestimates the negative association between computerization and unincorporated SE share growth. Regarding the validity of our shift-share IV, the instrument for computerization proves to be strong, with an F-test 174.983 compared to the critical 10% value in Stock and Yogo (2005) of 16.38. The instrument also passes the conventional under-identification test at the 1% significance level. Regarding the economic significance of the IV estimate, a one percent increase in computerization leads to a 0.791 percent decrease in the share of unincorporated SE in ten years. This estimate can be extrapolated to say that a ten percent increase in computerization can lead to a 7.9 percent reduction in the unincorporated SE share, which could have a substantial impact. This is our main IV estimate finding.

It is possible that during our sample period, the China syndrome as in Autor et al. (2013) is confounding the effects of computerization on the share of unincorporated SE. To ensure the impact we are capturing is due to computerization, we follow Autor et al. (2013) to use changes in Chinese imports by other high-income countries to instrument for the difference in import exposure per CZ worker. Column (9) shows the IV estimates when we control for the China syndrome. The coeffi-

[26]The imports are apportioned to the region according to its share of national industry employment, see Autor et al. (2013) for details.

cient for computerization is -0.737, which is quantitatively comparable to -0.791 from our baseline IV estimate in column (8). The coefficient for the import competition is statistically and economically insignificant within our framework. We also report the statistics of the instruments' under- and weak-identification test, and the instruments are still valid. In all, the "China shock" plays a minor role when computerization is considered in explaining the decline of unincorporated SE.

We also address the possibility that geographical distribution can affect the computerization rate and the unincorporated share differently. Hence, we consider the heterogeneous spatial impact of computerization in metropolitan versus non-metropolitan areas with interactions. A metropolitan dummy defines whether a city exists within the commuting zone. Column (10) shows the coefficients for the computerization, metropolitan, and their interaction term. As a standalone effect, computerization reduces the unincorporated SE share, whereas being in the metro area increases the share. The coefficient is statistically significant and negative as an interaction of the two. Since the level of computerization is generally higher in metropolitan areas, we expect the negative impact to be more substantial. The interaction term confirms this pattern: a one percentage point increase in the size of a metropolitan area leads to a 0.304 percentage point decline in the unincorporated SE share relative to that in non-metropolitan areas. This finding is also consistent with the efficiency-augmenting hypothesis because incorporated firms cluster in metropolitan areas. Therefore, unincorporated SE in urban areas face more competition than non-metropolitan areas.

Table 6: Computerization and the Change in the Unincorporated SE Share by Decade

	1990-2000		2000-2010		
	$2SLS_{25pct}$	$2SLS_{100pct}$	$2SLS_{25pct}$	$2SLS_{100pct}$	$2SLS_{3y}$
	(1)	(2)	(3)	(4)	(5)
Computerization	-0.314*	-0.326*	-1.067***	-1.046***	-0.680***
	(0.168)	(0.185)	(0.237)	(0.238)	(0.242)
Constant	0.128*	0.132*	0.459***	0.448***	0.138
	(0.068)	(0.073)	(0.137)	(0.138)	(0.135)
Observations	722	722	722	722	722
R^2	0.391	0.391	0.370	0.370	0.381

Notes: All the estimations models control for state and time-fixed effects and are weighted by the share of each commuting zone in the national population in each sample year, and the standard error is clustered at the commuting zone level. * Denotes significance at a 10% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.1.3 Subsample Decadal Analysis

The baseline estimates from Table 5 reveal a significant adverse impact of computerization on unincorporated SE. This finding underscores the crucial role of computerization in shaping the US economy between 1990 and 2010. To further examine this, we segment our sample by decade, providing separate estimates for 1990-2000 and 2000-2010. This approach not only enhances our understanding of the impact of computerization but also offers a unique perspective by potentially capturing the heterogeneous effects of recent advancements in computerization.

Firstly, dividing the sample by decade allows us to discern any evolution in the utilization of computers at work. The technological landscape in the 2000s may exhibit greater maturity compared to the 1990s, characterized by various updates and software programs. Secondly, our measure of computerization may serve as a proxy for the emergence of Web 2.0 during the late 2000s, coinciding with the nationwide proliferation of digital platforms, online order/payment systems, and the gig economy.[27] Table 6 presents the subsample decadal analysis estimates. For each decade, we also divide the sample by the 25 percent computerization and the 100 percent computerization at work. Columns (1) and (3) show that computerization's negative impact is more decisive in the second sample period. A similar pattern can be confirmed by columns (4) and (5) when we use computerization at 100% as the robustness measure.

Due to the Global Financial Crisis, 2010 saw a significant recession with an unusually high unemployment rate. Although the unincorporated SE share does not vary with the business cycle, we recalculate this share for 2010 using the 3-year average (2009, 2010, and 2011) to alleviate any concerns over temporary bias caused by the cyclical recession. Column (5) presents 2SLS estimates for the newly calculated 2010 unincorporated SE share. We find results that are smaller but statistically comparable to those in our baseline regression, which robustly confirms the negative relationship between computerization and the change in unincorporated SE over time. All columns underscore that the declining trend of the unincorporated SE share will likely continue as long as the efficiency-augmenting effect persists and is stronger than the restructuring effect.

5.2 Changes by Major Industry Groups

In Appendix B.2.3, we have demonstrated that the impact of computerization on unincorporated SE hinges on wage workers in industries with differing efficiency augmentation elasticity. As addressed in our hypotheses, high-computerization industries empower wage workers to counter restructuring effects and amplify efficiency gains from computerization, reducing unincorporated SE share. Conversely, low-computerization industries may heighten restructuring's impact on wage workers, limiting benefits and potentially increasing unincorporated SE.

Motivated by our conceptual framework, this section empirically tests the net impact of computerization across various industries. Specifically, we adopt the categorization provided by Census IPUMS,[28] which organizes industries into six broad groups. Our investigation revolves around scrutinizing the dynamic between computerization and unincorporated SE across industries characterized by differing levels of computerization. This analysis aims to present empirical findings that highlight the contrasting net effects of computerization in industries with high versus low levels of computerization. Furthermore, our inquiry extends to whether labor restructuring can account for the upsurge in the unincorporated SE share in industries with different computerization

[27] This encompasses the e-commerce landscape, characterized by the "long tail" phenomenon of niche products, as discussed in Kendall and Tsui (2011). These trends have evolved further into non-fungible tokens (NFTs), as explored in Borri et al. (2022).

[28] We exclude the agricultural sector in this analysis.

levels.

5.2.1 Estimating the Missing Unincorporated SE in High Computerization Industries due to the Efficiency-Augmenting Effect

Computerization augments the productivity and efficiency of wage workers, especially higher-skilled workers, which benefits incorporated firms through economies of scale. The efficiency-augmenting effect increases the comparative advantage of working in incorporated firms, especially those adopting a high level of computerization, over entering unincorporated SE. To show this argument, we group non-agricultural industries into six categories and estimate the impact of computerization on the low- and high-computerization groups. High-computerization industries experience more substantive efficiency gains than low-computerization industries and, therefore, have a higher efficiency-augmentation elasticity. If our hypotheses hold, we should see a more substantial decline in the unincorporated SE share in high-computerization industries.

The 2SLS estimation results are summarized in Table 7, Panel A. The first three columns of Table 7, Panel A, estimate the relationship between computerization and the changes in unincorporated SE share in the three industries with the lowest computerization. We repeat the same 2SLS estimation strategy and report the coefficients for industries with high computerization in the last three columns. The contrast between industries with low and high computerization is dramatic. Computerization appears to have a positive but insignificant impact on the change in the unincorporated SE share in all three industries with the lowest computerization. However, the impact of computerization on industries with high computerization is significantly negative. The negative impact is particularly pronounced for the manufacturing and wholesale industries. This finding is expected because the efficiency-augmenting effect is believed to be stronger in the high-computerization industries. In contrast, the restructuring effect can offset the impact on the low-computerization industries. The empirical findings align with the hypotheses put forth in the conceptual framework section.

To further show that the negative impact is driven by efficiency-augmenting gains attributable to computerization, we estimate wage workers' log hourly salary change in high- and low-computerization industries in Table 7, Panel B. The hourly wage is a standard proxy for labor's productive efficiency (Autor and Dorn; 2013). We find that the hourly wage rises significantly in industries with high computerization, such as the manufacturing and finance industries (columns (4) and (6) of Panel B). The unincorporated SEs are not growth-oriented, and unincorporated establishments are generally smaller than incorporated companies. Thus, it is more difficult for unincorporated SEs to compete with incorporated firms as the efficiency-augmenting effect becomes more substantive.

On the other hand, industries with low computerization are associated with a decline in real hourly wages. Columns (1)-(3) of Panel B show that a one percentage point increase in computerization in the CZ is associated with a 28.6 percent real hourly wage decrease in the in-person service industry, a 21.4 percent decrease in the construction/transportation industry, and a 48.3 percent decrease in the retail trade industry. This aligns with the restructuring effect, whereby computerization increases the unincorporated SE share in the in-person service industry. We explore this

Table 7: Computerization, the Change in the Unincorporated SE Share and the Change in Real Hourly Wage in the High and Low Computerization Industries

	I. Low Computerization Industry			II. High Computerization Industry		
	Personal service	Construction/ transportation/etc.	Retail trade	Manufacturing	Wholesale trade	Finance/ etc.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Δ in the share of unincorporated SE by industry						
Computerization	0.711	0.351	0.092	-2.300***	-3.604***	-0.476**
	(0.748)	(0.415)	(0.572)	(0.725)	(0.979)	(0.193)
Constant	-0.291	-0.299	-0.337	0.938**	1.882***	0.204*
	(0.448)	(0.228)	(0.310)	(0.398)	(0.550)	(0.111)
R^2	0.117	0.484	0.106	0.086	0.038	0.156
Panel B. Δ in the real hourly wage by industry						
Computerization	-0.286*	-0.214***	-0.483***	0.652***	-0.148	0.365***
	(0.168)	(0.080)	(0.111)	(0.149)	(0.186)	(0.096)
Constant	0.120	0.166***	0.269***	-0.268***	0.194*	-0.096*
	(0.100)	(0.045)	(0.062)	(0.081)	(0.101)	(0.052)
R^2	0.411	0.204	0.567	0.229	0.165	0.341

Notes: Each estimated coefficient in Panels A and B is based on a separate 2SLS regression with N = 1,444 (2 time periods $\times$ 722 commuting zones). All the estimations models control for state and time-fixed effects and are weighted by the share of each commuting zone in the national population in each sample year, and the standard error is clustered at the commuting zone level. The hourly wages are deflated by the national Personal Consumption Expenditure deflator. * Denotes significance at a 10% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

mechanism and present more evidence in Table 8.

5.2.2 Estimating the Growth of the Unincorporated SE in Low Computerization Industries due to Restructuring Effects

We turn now to the restructuring effect. In our context, this effect pertains to the scenario where wage workers, particularly those engaged in routine-intensive tasks within incorporated firms, are displaced by computerization/automation. Consequently, they transition into unincorporated SE.

This mechanism operates under the assumption that computerization primarily displaces less-educated, lower-skilled labor. However, tracking individual-level occupational selection changes over an extended panel (20 years) setting is constrained by data limitations. To address this challenge, we assume that newly displaced wage workers and other non-college-educated unincorporated SE individuals exhibit similar industrial entry preferences. Consequently, wage workers replaced by computerization, particularly those in routine-intensive occupations and lacking a college education, are likely to transition to industries preferred by unincorporated SE individuals. To validate this assumption, we first present the entry preferences of non-college-educated unincorporated SE individuals across various industries based on summary statistics (Panel A) in Table 8. Then, we follow it with a multinomial model (Panel B). Second, we estimate whether a higher

Table 8: Noncollege Unincorporated SE Labor Share and Industrial Preference, and the Effect of the Routine-intensive Occupations Share across Industries

	I. Low Computerization Industry			II. High Computerization Industry		
	Personal service (1)	Construction/ transport/etc. (2)	Retail trade (3)	Manufacturing (4)	Wholesale trade (5)	Finance/ etc. (6)
Panel A. Noncollege labor share in unincorporated SE						
Non-college	0.513	0.593	0.471	0.440	0.392	0.272
	(0.112)	(0.105)	(0.096)	(0.159)	(0.170)	(0.099)
Panel B. Disaggregated multinominal logit estimates: marginal effect on unincorporated SE industrial selection						
Non-college	0.170***	0.023***	0.003***	0.043***	-0.248***	0.010***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Panel C. 2SLS estimates: Δ in the share of the unincorporated SE						
Share of routine-intensive occupations	3.197***	-0.687	1.333**	-4.646***	-2.768	-0.323
	(1.122)	(0.799)	(0.579)	(1.153)	(2.164)	(0.408)
Constant	-0.788**	0.084	-0.657***	0.966***	0.674	0.032
	(0.351)	(0.229)	(0.164)	(0.323)	(0.609)	(0.117)
R^2	0.123	0.479	0.119	0.076	0.040	0.152
Observations	1,444	1,444	1,444	1,444	1,427	1,444

Notes: The summary statistics in Panel A are first calculated at the industry-CZ level and then weighted and aggregated at the industry level by the share of each commuting zone in the national population in each sample year. $Pseudo\ R^2$ and the number of observations for multinominal logit estimation are 0.964 and 742,695, respectively. Each estimated coefficient in Panel C is based on a separate 2SLS regression with N = 1,444 (2 time periods $\times$ 722 commuting zones). All the estimations models control for state and time-fixed effects and are weighted by the share of each commuting zone in the national population in each sample year, and the standard error is clustered at the commuting zone level. * Denotes significance at a 10% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

start-of-period routine-intensive occupational share leads to a rise in the unincorporated SE share in the preferred industries of non-college-educated, unincorporated SE individuals (Panel C).

Panel A provides summary statistics regarding the share of non-college-educated unincorporated SE individuals across different industry groups. A notable trend emerges: in industries characterized by low levels of computerization, such as in-person services (51.3 percent) and construction/transportation (59.3 percent), most unincorporated SE individuals do not possess a college degree. Conversely, in industries with high levels of computerization, such as finance, the proportion of non-college-educated individuals is substantially lower (27.2 percent). This observation underscores the alignment between the distribution of computerization and the ratio of non-college-educated unincorporated SE individuals, revealing significant variations across industries. These findings have important implications for understanding the impact of computerization on occupational selection and industry preferences.

To delve deeper into the industrial entry preferences of non-college-educated unincorporated SE individuals, we employ a multinomial logit model and present the marginal effect estimates in

Panel B. This analysis allows us to gauge the impact of educational attainment on the selection of unincorporated SE across various industries at the individual level. Unlike the CZ-level regressions, we restrict the sample to unincorporated SE individuals. The findings from Panel B align with the summary statistics, revealing that non-college-educated, unincorporated SE individuals are more inclined to enter the in-person service industry compared to industries characterized by high levels of computerization. This exercise, coupled with the summary statistics, underscores the preference of non-college-educated, unincorporated SE individuals for working in the in-person service industry.

In Panel C of Table 8, we test the restructuring effect hypothesis by regressing the routine-intensive occupation share within a commuting zone on the change in the unincorporated SE share by industry. The first three columns show that commuting zones with initially high routine-intensive occupational shares see an increase in low-computerization industries, particularly the in-person service and retail industries. For instance, a ten percentage point higher routine-intensive share, equal to the gap between commuting zones at the 80th ($RSH^{\mathrm{P20}} = 0.567$) and 20th ($RSH^{\mathrm{P20}} = 0.468$) percentile, predicts the unincorporated SE share growth that is approximately 32.0 log points higher in the in-person service industry between 1990 and 2010.

In contrast, the impact of the routine-intensive share on industries with high computerization presents a different picture. For the manufacturing industry, a higher routine-intensive share is associated with a substantial decline in the unincorporated SE share. In the wholesale trade and finance industries, which have the highest computerization among the industrial groups, the routine-intensive occupational share is not significantly correlated with the unincorporated SE share. Panels A and B show that non-college-educated labor is least likely to enter the wholesale and finance industries. As expected, the replaced labor force is less likely to select industries with high computerization.

Autor and Dorn (2013) argue that automation is one of the main factors contributing to the recent job market polarization. However, like the authors of many related studies (Nordhaus; 2007; Autor et al.; 2003; Acemoglu and Autor; 2011), we note that the unincorporated SE has largely been ignored in discussions concerning the job market polarization. Computerization does not polarize the unincorporated SE; it relocates unincorporated labor towards only low-technology-augmented sectors. However, because unincorporated SE entry has a clear industrial selection pattern, the restructuring effect does not lead to a uniform change in the unincorporated SE share across industries.

5.3 Alternative Hypothesis

As shown earlier in Figure 2, the unincorporated SE share is not uniformly distributed across industries. In addition, low-skill service and abstract-intensive occupations have outgrown routine-intensive occupations since 1980 in the United States. Hence, we consider two alternative hypotheses. First, we explore the hypothesis that computerization-induced industrial compositional change is sufficient to explain the decline in the unincorporated SE share. Second, we turn to the alternative income effect hypothesis to explain the positive relationship between computerization and the unincorporated SE share in industries with low computerization, particularly in the in-person service industry. The income effect refers to rising income at the top of the wage distribution, stimulating demand for in-person services among wealthy households. To determine this, we test whether the rising top quantile wage increases the unincorporated SE share. Overall, we conclude neither the industrial compositional change nor the income effects hypothesis is the main factor contributing to the empirical change in the unincorporated SE share.

5.3.1 Local Industry Compositional Change

Recall that the computerization measure is based on the local commuting zone's industrial composition and the national average industrial computer usage rate. The simulated computerization is calculated as the local CZ industrial composition in 1980 multiplied by the national industrial computerization in 2000. We fixed the industrial composition in 1980, the year we typically assume no computerization can be observed. Doing so allows only computerization variation from within industry growth and holds industrial composition constant within the sample periods. Appendix D.2 further discusses the association between the original and simulated measures of computerization. Suppose the industrial compositional change is sufficient to explain the contraction in the unincorporated SE. In that case, the results inferred from the survey-based data and simulated computerization levels should be substantively different.

Table 9: Simulated Computerization and the Change in the Unincorporated SE Share by Decade

	OLS		2SLS	
	1990-2000	2000-2010	1990-2000	2000-2010
	(1)	(2)	(3)	(4)
$\text{Computerization}_{\text{simulated}}$	-0.199	-1.121***	-0.361*	-1.274***
	(0.133)	(0.216)	(0.195)	(0.298)
Constant	0.083	0.474***	0.144*	0.555***
	(0.058)	(0.125)	(0.076)	(0.165)
Observations	722	722	722	722
R^2	0.390	0.375	0.388	0.374

Notes: Each estimated coefficient is based on a separate decade regression with N = 722 (1 time period × 722 commuting zones). All the estimations models control for state and time-fixed effects and are weighted by the share of each commuting zone in the national population in each sample year, and the standard error is clustered at the commuting zone level. The simulated computerization is calculated by fixing the industrial mix in 1980 and only allowing the variation to come from the within-industry computerization growth. * Denotes significance at a 10% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To verify whether industrial compositional change affects our conclusions, the left and right panels of Table 9 report the OLS and 2SLS estimations of the effect of the simulated computerization on the change in the unincorporated SE share. Consistent with the computerization measure reported in Table 6, the second period's negative impact is more significant. The estimated coefficients in the 2SLS regressions are all slightly smaller than their counterparts in columns (4) and

(5) of Table 6 because the simulated computerization measure ignores the fact that industrial composition grows towards the computer-intensive industries. Nevertheless, the results show that the impact of the simulated computerization measure is statistically comparable to the estimates based on survey-based data computerization. Therefore, industrial compositional change is only one of the factors contributing to the decline in the unincorporated SE share.

5.3.2 Income Effect Hypothesis

Our second alternative hypothesis is the income effect, which refers to the rising income at the top of the wage distribution, stimulating demand for in-person services among wealthy households. To determine this, we test whether the rising top quantile wage increases the unincorporated SE share, especially in the in-person service industry.

In Appendix D.3, we first estimate whether computerization is positively associated with top-quantile wage growth. The estimated quantile coefficients confirm a highly heterogeneous effect of computerization on the wage for employed labor. The top quantile's hourly wage increases substantially due to a higher computerization. To explore the possible link between top-quantile wage growth and the change in the unincorporated SE share, we include hourly wage change at the 95th percentile in the estimation. However, the result in Table 10 does not support the income hypothesis. Specifically, in column (1) of Table 10, the estimated coefficient shows that the hourly wage change at the 95th percentile has no significant effect on the difference in the unincorporated SE share. We include the hourly wage change at the 95th percentile and computerization and report the estimated results in column (2). The 95th percentile wage change is also insignificant when the computerization measure is included. For columns (3) and (4), we repeat the same estimation strategy and use the 90th percentile hourly wage change to check the robustness of the conclusion. Consistent with the conclusion based on wage growth at the 95th percentile, wage growth at the 90th percentile has no statistically significant positive impact on the unincorporated SE share.

Although the rise in wages in the top quantiles cannot predict an increase in the unincorporated share on average, it is reasonable to assume that it may work as a demand shifter and positively affect the in-person service industry. The right panel (columns (5)-(8)) of Table 10 explores the potential contribution of the income effect to the rising unincorporated SE in the service industry by augmenting the baseline regression model with the additional top quantile wage growth measures. In columns (5) and (6), we use changes at the 95th percentile of the log weekly wage distribution among full-time, and full-year workers in the CZ to capture wage structure shifts generated by income effects. In columns (7) and (8), we also report the results of the robustness exercise using the 90th percentile of the log weekly wage distribution.

Nevertheless, these proxies for the income effect do not have a substantial direct relationship with the change in the unincorporated SE share. A rise in the 95th/90th percentile of hourly wages is weakly correlated with a declining unincorporated SE share when computerization is included, and this relationship becomes insignificant when we include only the income proxy measures. Therefore, we conclude that the income effect cannot explain the growth in the unincorporated low-skill service or in-person industries.

Table 10: The Change in Unincorporated SE share by Income

	$10 \times$ *Annual change in share of unincorporated SE*				$10 \times$ *Annual change in the share of unincorporated SE in the service industry*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Computerization		-0.791***		-0.746***		1.022		1.159
		(0.164)		(0.174)		(0.765)		(0.796)
Δ ln(P95) hourly wage	-0.114	-0.000			-0.378	-0.525		
	(0.084)	(0.089)			(0.319)	(0.333)		
Δ ln(P90) hourly wage			-0.239**	-0.072			-0.462	-0.720*
			(0.096)	(0.110)			(0.383)	(0.418)
Observations	1,444	1,444	1,444	1,444	1,444	1,444	1,444	1,444
R^2	0.153	0.167	0.157	0.168	0.116	0.120	0.116	0.121

Notes: Each estimated coefficient is based on a 2SLS regression with N = 1,444 (2 time periods $\times$ 722 commuting zones). All the estimations models control for state and time-fixed effects and are weighted by the share of each commuting zone in the national population in each sample year, and the standard error is clustered at the commuting zone level. * Denotes significance at a 10% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.4 Discussions

Our conceptual framework posits that technological changes can lead to ambiguous outcomes. On the one hand, computerization helps with the burgeoning of small services and digital products, as indicated by our restructuring effect hypothesis. On the other hand, as documented in Kuratko and Audretsch (2022), a few tech giants are replacing smaller entrepreneurs, which creates more productive firms, as indicated by our efficiency-augmenting effect hypothesis. Therefore, we utilize empirical evidence from the early adoption of computers in the 1990s to document their overall impact on the US economy. Based on our findings from the empirical analysis, we designate this section to discuss the relevance of our study to the literature on entrepreneurial human capital and emerging technological innovation.

For some, it can be construed that the digital divide may further polarize the economy between incorporated and unincorporated SE. Acemoglu and Restrepo (2020) find that one more robot per thousand workers reduces the employment share by 0.2 percentage points. Our finding is consistent with the literature's finding of the negative association between technological advancements and employment. We show that this relationship is only apparent for the unincorporated entrepreneurs by a 0.79 percent decrease. Moreover, we find an extra 0.304 percentage point reduction in the unincorporated entrepreneurs for metropolitan areas.

However, we view our findings as more nuanced and multi-dimensional, which should be interpreted as a positive development: computerization enables the more productive firms to scale up and utilize a larger fraction of the resources in the economy. In this sense, the economy is becoming more dynamic and productive. While our sample period has focused on the 1990-2010 era, the same principle can be applied to other technological developments in the later eras, mainly the rise of Web 2.0 since the early 2010s and the emergence of AI since the 2020s. Burtch et al. (2018) study the implications of the gig economy and Web 2.0, such as Uber X, on entrepreneurial activities, measured by the rate of campaign launches at Kickstarter, a crowd-funding platform. They document that the rollout of Uber X in the local area predominantly reduces unincorporated entrepreneurial ventures.[29] Brynjolfsson et al. (2023) offer an analysis of generative AI at work for customer support agents and find that the new technology increases productivity (issues resolved per hour) by 14 percent on average. The increase in productivity is more significant for low-skilled workers, at 34 percent.

In all, we find a common theme among our paper (computerization), Burtch et al. (2018) (gig economy), and Brynjolfsson et al. (2023) (AI): technological change can improve productivity, with heterogeneous effects across workers. Given that emerging tools can change the way workers perform and learn, it is vital to understand the implications of entrepreneurial human capital on the US economy. Ehrlich et al. (2017) predicate that the investment in entrepreneurial human capital at the average firm-level positively affects productivity growth.[30] Given the common theme of any skill-biased technological innovation's (e.g., computers, digital platforms, AI) impact in increasing the share of innovative entrepreneurs in incorporated firms, we can reach a compatible inference that greater investment in research and development would lead to higher productivity growth (or economic efficiency level in our model).

6 Conclusion

From 1990 to 2010, the workplace witnessed a significant surge in computer adoption, with a 40.5 percent increase in computerization. Our study examines the profound implications of this computerization on local labor markets in the United States, with a specific focus on its impact on unincorporated SE. Over the same period, we observed a stark decline of more than 25 percent in the percentage of individuals engaging in unincorporated SE. Our empirical findings strongly suggest that the proliferation of computerization played a pivotal role in driving this overall decline. This raises significant concerns about the future landscape of employment in the United States, particularly as emerging technologies continually reshape the labor market.

To systematically analyze the multifaceted effects of computerization, we have outlined a conceptual framework that encapsulates both its positive and negative implications. On the one hand, computerization can foster unincorporated SE, a phenomenon we term the restructuring effect. Conversely, it can also diminish the prevalence of unincorporated SE, a dynamic referred to as the efficiency-augmenting effect.

The restructuring effect anticipates a positive correlation between computerization and changes in the share of unincorporated SE, particularly in industries with low levels of computerization, such as those in the in-person service sector. Conversely, the efficiency-augmenting effect underscores the productivity gains derived from computerization among wage workers. As comput-

[29] They also provide an excellent literature review regarding the effect of the gig economy on how workers develop new employment opportunities.

[30] Subsequently, studies like Qin and Kong (2021), Cintio (2022), and Sima (2023) signify the importance of entrepreneurial human capital.

erization progresses, it complements the productivity of incorporated firms, with wage workers assuming roles centered on abstract thinking, creativity, problem-solving, and coordination. This trend, coupled with the increase in entry and fixed costs for firms, facilitates the relocation of resources, rendering unincorporated SE individuals less competitive compared to their counterparts in incorporated firms within the same industry. In contrast to the employment and wage polarization experienced by wage workers, individuals engaged in unincorporated SE find themselves marginalized toward the lower end of the income spectrum.

Our empirical analysis aims to ascertain which effect, as per our hypotheses, has exerted a dominant influence on the US economy since the advent of computerization in the 1990s. Leveraging recent advancements in the field, we employ Bartik (or shift-share) instruments as the primary empirical framework, incorporating several IV tests proposed in Borusyak et al. (2022). Our estimates reveal that a one percent increase in computerization correlates with a 0.79 percent reduction in the share of unincorporated SE. This effect is notably more pronounced in metropolitan areas, where computerization correlates with an additional 0.304 percentage point decrease. IV estimates lend further credence to the dominance of the efficiency-augmenting effect, which has contributed significantly to the overall decline in unincorporated SE.

Distinguishing our study from previous research that predominantly examines entrepreneurship in general, we focus on unincorporated SE characterized by low capital intensity, investments, and average sales. Consequently, while recent findings support entrepreneurship growth driven by computerization, our empirical conclusions offer a more nuanced perspective. Importantly, these findings remain robust across various alternative settings, including adjustments for the impact of the China shock during the sample period.

We envision our paper as a blueprint for investigating the nexus between new technologies and their implications on the labor market. While our analysis centers on computerization, the framework readily adapts to emerging technologies such as robotics, augmented reality, and artificial intelligence. Furthermore, we underscore the importance of employing shift-share instruments and rigorous validity and robustness tests in designing empirical studies in this domain. Finally, a deeper examination of the factors driving the decline in entrepreneurship, particularly among unincorporated SE individuals, holds significant policy implications and promises to enhance our understanding of economic dynamism.

A Supplementary Figures and Tables to Section 2

A.1 Incorporated SE vs. Unincorporated SE

Table A.1: Comparison between Incorporated and Unincorporated SE (1990-2010)

	Incorporated SE	Unincorporated SE
Education		
$<$ High school	0.061	0.154
High school	0.223	0.296
Some college	0.260	0.265
College	0.278	0.177
$>$ College	0.178	0.108
Firm Size		
<10	0.597	0.769
10-99	0.281	0.140
100-499	0.039	0.016
500+	0.068	0.038
Income		
Total income ($1,000)	69.760	37.822
Industrial Sectors		
Personal Service	0.031	0.070
Construction/transport/etc.	0.197	0.211
Retail trade	0.183	0.130
Manufacturing	0.078	0.038
Wholesale trade	0.066	0.026
Finance	0.401	0.411

Notes: The values in *education* represent the proportions of self-employed labor within each education category. Meanwhile, the 'firm size' variable in CPS is categorized into broader groups, indicating the number of employees working for the respondent. The values in *firm size* denote the proportions of self-employed labor within each firm size category. Total income is adjusted for inflation using the IPUMS variable 'CPI99,' thereby converting the dollar amounts to their equivalent values in 1999. This adjustment ensures consistency with the 2000 CPS data and facilitates accurate comparisons over time.

A.2 Why Remove Unpaid Family Workers from the Sample?

The exclusion of unpaid family workers from our analysis is justified by the following substantial differences, which can potentially skew or bias the results if they were included. This approach ensures a more accurate and relevant analysis of the labor market as it pertains to our study's focus.

1. Nature of Work Involvement: Unpaid family workers typically engage in non-market transactions, contributing to family-run businesses or household activities without receiving formal wages. Their involvement is often not a part of standard market transactions, which challenges their assessment using conventional wage and labor models.

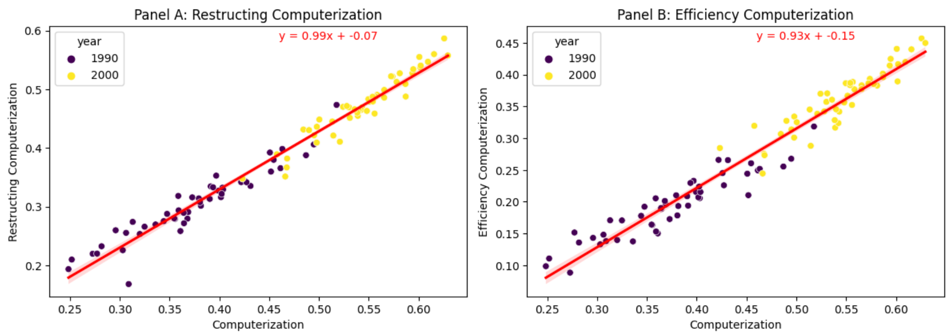
2. Valuation of Labor: Assigning a monetary value to the labor of unpaid family workers is complex due to the lack of a direct salary or hourly wage. This complication contrasts with traditional employees who receive clear, quantifiable compensation, making it difficult to analyze wage levels and structures within the context of unpaid family work.

3. Work Hours Variability: The working hours of unpaid family workers tend to be irregular and highly variable, diverging significantly from the more structured and consistent working hours of regular wage earners. This variability poses challenges when attempting to draw comparisons between these two groups.

4. Diverse Motivations: The motivations for engaging in unpaid family work are often distinctly different from those in paid employment. Influenced by family obligations, cultural norms, or personal choices rather than market forces, these motivations can significantly affect the nature and intensity of their labor. This difference renders their work not easily comparable to typical wage-earning scenarios.

A.3 Defining Computerization within the Conceptual Framework

Figure A.1: Computerization within Our Conceptual Framework



Notes: Restructuring and Efficiency refer to the two channels of our conceptual framework in Section 3.

In our conceptual framework (Section 3), we discuss two channels: the restructuring effect and the efficiency-augmenting effect, both of which are occupation-related. The restructuring effect posits that computerization leads to the displacement of certain tasks previously performed by wage employees, while the efficiency-augmenting effect suggests that computerization complements the skills of wage employees working in incorporated firms. Therefore, a survey primarily focused on the occupation of a particular worker can provide insightful evidence for both channels.

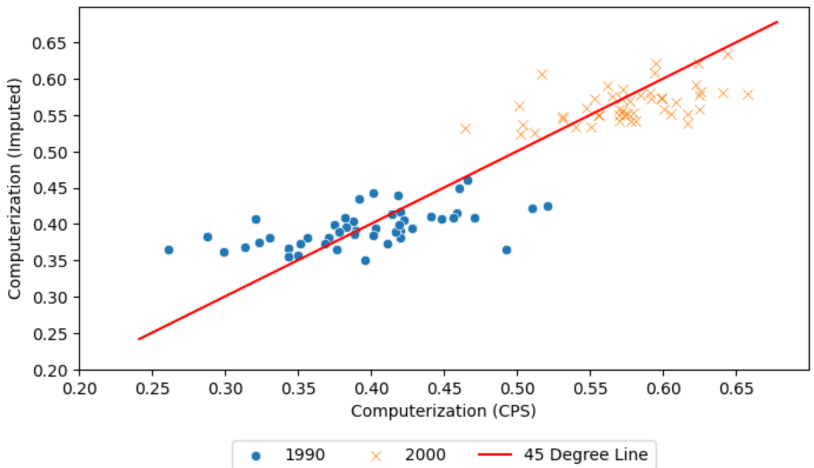
To ensure that our definition of computerization aligns with these two channels, we carefully measure computerization to reflect their merits. The computer and internet use supplement survey offers valuable insights in this regard. For example, in the 1989 CPS October Supplement, the survey not only asks whether a computer is used at work but also inquires about specific purposes for which it is used. These purposes encompass a wide range of tasks, including word processing, bookkeeping, computer-assisted design, and more, providing a comprehensive perspective on the diverse applications of computers in the workplace.

We categorize the purposes of using computers into two groups to reflect the restructuring and efficiency-augmenting intentions. The restructuring channel includes purposes such as bookkeeping, calendar/scheduling, communications, desktop publishing/newsletters, electronic mail, graphics, inventory control, invoicing, sales, and word processing. On the other hand, the efficiency-augmenting channel includes purposes such as analysis, databases, programming, and spreadsheets. It is important to note that these purposes are not mutually exclusive.

The left panel of Figure A.1 displays the correlation between computerization and the restructuring purpose of computer use. In contrast, the right panel depicts the correlation between computerization and the efficiency-augmenting purpose of computer use. In both cases, our measure of computerization exhibits a highly significant positive correlation with both the restructuring and efficiency-augmenting purposes of using computers.

A.4 Survey-based data vs Imputed Computerization

Figure A.2: Survey-based data vs Imputed Computerization



Notes: Computerization (CPS) refers to the survey-based data. Computerization (Imputed) refers to the constructed computerization variable in a commuting zone.

A.5 Computerization Across States

Figure A.3: Computerization Difference by US State

Notes: This figure plots the computerization difference between 1990 and 2000 by state (left panel) and the computerization by unincorporated versus incorporated by state (right panel).

To understand the heterogeneous nature of computerization in the United States, Appendix Figure A.3 depicts different rates of computerization by state. The left panel exhibits the difference in computerization between the incorporated and unincorporated SE. The y-axis represents the difference in 1990, and the x-axis represents the difference in 2000. The comparison between the decades will indicate whether computerization changes over time. For instance, Washington state had almost zero difference in computerization between the incorporated and unincorporated SE in 1990. In contrast, the difference is nearly 40 percent in 2000, revealing a significant computerization development over a decade. We draw the 45-degree line to discern state-by-state changes in the difference in computerization between the incorporated and unincorporated SE from 1990 to 2000. Computerization difference in Nevada, for instance, does not change over the decade and remains constant at 30%. Notable states on the left of the 45-degree line are Arkansas, West Virginia, and Missouri, as they had over 40 percent difference in 1990. Still, the difference is between -10 percent and slightly over 10 percent.

To examine the pronounced improvement, the right panel of the figure depicts the pooled average of computerization between 1990 and 2000 between the incorporated and unincorporated SE. The difference between the two groups is quite astonishing. Except for Maryland, the incorporated self-employed have a higher computerization than the unincorporated. This result indicates that the incorporated entities and the unincorporated SE have substantially different likelihoods of adopting computers in their workplaces, with this technology gap widening over time.

In our analysis of the sample period spanning 1990-2010, characterized by the nascent stages of internet development,[31] computerization is particularly pronounced for incorporated firms. At the same time, its impact is relatively constrained for unincorporated SE, such as those operated by unincorporated SE individuals. This discrepancy can be attributed to several factors, including the larger size, a division of labor encompassing different types of skilled wage employees, better internet accessibility, and greater availability of resources in incorporated firms. These factors collectively contribute to the more effective adoption of computer technology in incorporated firms. Our argument finds additional support in the evidence presented in the right panel of Figure A.3, which illustrates the disparity in computerization between unincorporated and incorporated firms during the 1990s and 2000s. According to this argument, the influence mechanism shows that some self-employed individuals seek opportunities to work for incorporated firms, causing a decrease in the number of individuals in unincorporated SE relative to the number of wage workers in incorporated firms.

[31] According to survey data from Statista Research Department, the adoption rate of the internet in the US was 18% in 1997, significantly lower than the rate of 85.5% recorded in 2020.

B Supplementary Information to Section 3

B.1 Conceptual Framework's Relation to the Literature

The distinctive pattern between the incorporated and unincorporated SE in relation to technology discusses the ability distribution in the labor market. As such, our study is closely related to the literature exploring the relationship between computerization and the low-skill wage worker's occupational choice. Autor and Dorn (2013) argue that computerization in recent decades has caused the cost of automating routines to fall, which leads to greater adoption of information technology and the reallocation of low-skill workers from routine task-intensive occupations into service occupations. These forces increased the demand for low- and high-skill service occupations in the United States from 1980 to 2005. Similar findings regarding labor market polarization are pervasive in tests with different specifications (Autor et al.; 2006, 2008) and data from advanced economies such as the United Kingdom (Goos and Manning; 2007) and European countries (Goos et al.; 2009; Michaels et al.; 2014).

In recent decades, many studies have offered hypotheses and frameworks to explain employment/job polarization. The canonical model is based on the skill-biased technological change hypothesis (Tinbergen; 1974; Autor et al.; 2006, 1998; Welch; 1973; Acemoglu; 2002), which suggests that improvements in technology increase the demand for more skilled workers because new technologies are more complementary to high-skill workers.[32] This popular hypothesis explains the increase in employment in high-skill occupations but is silent on the increased demand for low-skill occupations that appears in recent empirical evidence.[33]

To address the shortcomings of the canonical model, Acemoglu and Autor (2011) develop a task-based framework to analyze recent changes in employment distribution.[34] Unlike the canonical model, which imposes a one-to-one correspondence between skills and tasks (occupations), the task-based framework assumes that allocating skills to tasks is endogenous and responds to changes in technology involving substituting machines for routine tasks previously performed by labor. The task-based framework links employment polarization to the routine-biased technological change hypothesis (Goos et al.; 2014), or simply the routinization hypothesis (Autor et al.; 2003; Acemoglu and Autor; 2011). Specifically, computerization reduces the actual cost of performing standardized processing (Nordhaus; 2007), further changing the labor supplies of skills.

However, several recent studies have re-examined the labor market using alternative approaches

[32] Readers are referred to Goldin and Katz (2009); Acemoglu and Autor (2011) for overviews of the canonical model. The skill-biased technological change hypothesis is supported by rich empirical studies: (e.g., Carneiro and Lee; 2011; Autor et al.; 1998; Card and Lemieux; 2001; Atkinson; 2007)

[33] Acemoglu and Autor (2011) document several key trends the canonical model fails to interpret. First, it does not explain why less-educated workers have experienced real earnings declines in recent decades. Second, it does not reflect the disproportionate growth of employment in high-education, high-wage occupations and low-education, low-wage service occupations. Third, it does not allow computers and robotics to replace workers in routine-intensive occupations. Fourth, it assumes that technical changes are exogenous and do not respond to changes in labor market conditions. Finally, it is silent on how offshoring affects labor market inequality.

[34] Other, less flexible or tractable task-based modes of the impact of technology and outsourcing on the labor market include those of Feenstra and Hanson (1999); Acemoglu et al. (2015); Spitz-Oener (2006); Goos and Manning (2007); Grossman and Rossi-Hansberg (2008); Dorn et al. (2009); Autor and Dorn (2010); Costinot and Vogel (2010).

and have found evidence contradicting employment polarization, challenging the narrative of routinization. By tackling the issue of systematic occupational mismatch, Lefter et al. (2011) discovered compelling evidence supporting a persistent trend of employment growth in high-skill jobs and a decline in certain middle-skill jobs, with no significant variations observed between the 1980s and 1990s. Mishel et al. (2013) found that the task-based framework fails to explain the most significant developments in wage trends observed since the end of 1970. Similarly, Hunt and Nunn (2022) discovered significant empirical weaknesses in the occupation-based approach adopted by previous literature. Instead, employing a worker-based approach, they found evidence of a decline in the proportion of workers earning middle wages since 1973, aligning with the findings of occupation-based analyses. However, they also observed a substantial increase in the proportion of high-paid workers without a corresponding rise in the share of low-paid workers. This contradicts the argument of employment polarization.

Despite some empirical evidence showing deviations from the modeling of the task-based framework regarding employment polarization, researchers still recognize its elegance and appreciate its richness compared to the canonical model it seeks to replace (Mishel et al.; 2013). Moreover, in contrast to previous studies that primarily focus on wage employees across three skill sets (low, middle, and high) and the phenomenon of employment polarization, our study emphasizes the utilization of the task-based framework to examine the decision between self-employment and wage employment. The empirical examination of applying the task-based framework to the decision of unincorporated SE has not been explored in the existing literature.

Within our framework, the task-based framework exhibits two simultaneous effects. First, it creates an economic incentive for employers to substitute workers with computers. Computerization entails greater automation and the utilization of computers to replace wage employees working in incorporated companies. Workers engaged in routine-oriented tasks are particularly susceptible to displacement through computerization. As a result, computerization leads to unemployment across various occupations, leading to many displaced workers transitioning to unincorporated SE for new opportunities. This phenomenon is referred to as the restructuring effect in this study.

Second, computerization complements workers and improves their productivity and efficiency. We term the second effect as the efficiency-augmenting effect. According to Autor et al. (2003), the efficiency-augmenting effect[35] particularly benefits non-routine task workers. Non-routine tasks encompass both abstract tasks and manual tasks. Since these two task types exist at opposite ends of the occupational skill spectrum, the task-based framework predicts a polarization of employment opportunities.[36] The efficiency-augmenting effect is also frequently mentioned in a large strand of literature studying firm size and information technology. From Brynjolfsson and Hitt (2000), computerization or investment in information technology enables transformation in organization process and structure by the decentralization/centralization of decision-making, autonomy, firm

[35] Acemoglu and Restrepo (2019) refers to it as the productivity effect and the reinstatement effect.

[36] Another popular non-technology-oriented hypothesis explaining job polarization in advanced economies points to the change in the structure of international trade; this is the offshoring hypothesis. As noted by Acemoglu and Autor (2011), the offshoring hypothesis can also be tested with the task-based model they developed, with specific tasks performed by medium-skill workers here performed by lower-paid foreign workers instead of machines. Blinder and Krueger (2013) also argue that, while conceptually different, offshorability is related to occupation routineness.

size, and coordination and communication in working processes, which leads to better business performance. Most empirical studies have also reported a significant positive effect of computerization on productivity (Byrd and Davidson; 2003; De Guinea and Webster; 2013; Hou; 2013; Sun; 2017; Marcolin et al.; 2018).

B.2 A Simple Model

B.2.1 Endogenous Skill Demand and Supply

We present a stylized model derived and expanded from Acemoglu and Autor (2011), showcasing the influence of computerization on the changing share of unincorporated SE. Our approach to identification utilizes the microeconomics choice decision model. We acknowledge that the model we have crafted serves as an illustrative tool and comes with its own set of limitations, as it excludes various pertinent factors such as trade, migration, endogenous skill formation, and capital accumulation that could influence unincorporated SE. The more comprehensive models put forth by Acemoglu and Restrepo (2020), Salgado (2020), Kozeniauskas (2022), and Jiang and Sohail (2023) offer a more detailed framework for understanding how macroeconomic factors influence the trajectory of entrepreneurship. These models should be regarded as more holistic theoretical frameworks than ours. The objective of our model here is to analyze the effects of computerization on the share of unincorporated SE while controlling for other factors.

Despite excluding several macroeconomic factors, the microeconomics-oriented model provides a picture of how computerization affects unincorporated SE through two main channels. In the empirical section, we include conditions not discussed in the model section and find that these two channels resulting from computerization remain robust in spite of these conditions.[37]

To allow an endogenous choice of both skill supply and demand in the model, we refer to the task-based framework of the labor market developed by Acemoglu and Autor (2011). Deviating from the conventional assumption of inelastic labor supply, we assume that the decision between starting an unincorporated SE and pursuing wage employment is endogenous. Furthermore, we assume that transitioning between career paths does not involve substantial learning or entry costs. Although recent studies (Lefter et al.; 2011; Mishel et al.; 2013; Hunt and Nunn; 2022) have presented empirical challenges to the notion of employment polarization within the task-based framework, the endogenous skill supply setting still offers a more realistic environment compared to the conventional canonical model. Additionally, the restructuring and efficiency-augmenting effects in this framework are intuitively straightforward when applied to the decision-making process of self-employment.

Endogenous skill supply In the endogenous skill-supply environment, we assume that a representative agent, denoted as i, has the choice to either pursue unincorporated SE or engage in wage

[37]We conducted several robustness tests, including examining the effects of offshoring and macroeconomic conditions. Although these conditions do have some impact on unincorporated SE, the two channels of computerization (restructuring and efficiency-augmenting effects) remain significant. While we cannot dismiss the influence of other factors, the two channels resulting from computerization are likely to be major contributors to the changing share in unincorporated SE.

work within unincorporated or incorporated business settings. This decision hinges upon the individual's self-employment/entrepreneur ability ($se(i)$) and employment ability ($e(i)$). We assume that each individual has one unit of time that can be allocated either to unincorporated SE or to working as a wage employee for other unincorporated or corporate business entities. Furthermore, we rank each individual's ability over the interval (0,1) such that individuals with a greater comparative advantage in employment ability relative to self-employment ability are classified closer to 0 and those who have a greater comparative advantage in self-employment ability are closer to 1 in the interval. More precisely, $e(i)/se(i)$ is strictly decreasing in i such that $lim_{i\to 0}e(i)/se(i) = \infty$ and $lim_{i\to 1}e(i)/se(i) = 0$. Given the assumption above, each individual allocates his or her time to either becoming an unincorporated SE or working as a wage worker. Individuals with $\pi se(i) > we(i)$ choose to open unincorporated SE and earn profit rate π, while individuals with $\pi se(i) < we(i)$ work as employees with wage rate w. Finally, no-arbitrage condition holds for individual $i^\star$ when

$$\frac{se(i^\star)}{e(i^\star)} = \frac{w}{\pi}.$$

The relative supply curve of wage employees (E) to unincorporated SE individuals (SE) is therefore

$$(\frac{E}{SE})_{supply} = \frac{\int_0^{i^\star} e(i)di}{\int_{i^\star}^1 se(i)di}, \tag{B.1}$$

where $0 < i^\star < 1$ is the threshold value of an individual $i^\star$ in interval $(0, 1)$. Any individual $i < i^\star$ chooses to be a wage employee, while an individual with $i > i^\star$ chooses to become an unincorporated SE. Furthermore, the relative supply curve E/SE is *upward sloping* given that $i^\star$ is strictly increasing with the increase in payoff ratio w/π.

One caveat of our setup is that we prioritize ranking the comparative advantage of an individual's employment ability over their self-employment ability. However, unlike the specification by Kozeniauskas (2022), we do not differentiate between wage workers as high- or low-skilled. This is because both high- and low-skilled workers have the potential to begin their own unincorporated SE businesses. To accommodate this aspect in our model, we refrain from making such distinctions among wage workers. An inherent limitation of this approach is that we cannot ascertain whether high-skilled or low-skilled workers are more likely to transition into or exit unincorporated SE, as demonstrated by Kozeniauskas (2022). Another simplification in the model is that we do not explicitly specify whether wage workers will be employed by unincorporated SE ventures, non-entrepreneurial firms, or any other type of business entity. In the model, they are all classified simply as wage workers. Consequently, the model will only be able to demonstrate changes in unincorporated SE, without the capability to discern whether employment in unincorporated SE increases or decreases with computerization.

Endogenous skill demand We now turn to the endogenous setting of skill demand. In a closed economy, we assume that a production/service section Y is combined with a continuum of different tasks ranging from non-routine to highly routined across a unit interval[38] such that

$$\ln Y = \int_0^1 \ln y(j) dj, \tag{B.2}$$

where $y(j)$ is the production level of task j. We assume that all markets are competitive and that the price of the final good is the numeraire. Non-routine jobs typically demand human creativity, adaptability, and problem-solving skills. Professions such as carpentry, gardening, plumbing, artistic endeavors, writing, and entrepreneurship exemplify this type of work. Often managed individually or with minimal hired labor, non-routine tasks are primarily carried out within unincorporated SE arrangements. This aligns with the 2014 CPS data, where only 13% of unincorporated SE individuals had at least one paid employee. On average, unincorporated SE ventures employed 5.6 paid workers if they hired staff.

On the other hand, highly routinized tasks, which involve performing repetitive and standardized actions efficiently, are more commonly associated with larger-scale non-entrepreneurial firm settings. This distinction is based on the higher computerization observed in the incorporated setting during the period of 1990-2010, as discussed and illustrated in Figure A.3. However, the skill demanded and tasks performed do not necessarily correspond one to one due to the endogenous setting of skill demand. Each task $y(j)$ follows a perfect substitution production function combined with the aggregation of individuals who endogenously choose either unincorporated SE (SE) or wage employment (E), as well as computer capital (K) such that

$$y(j) = A_{SE} AP_{se}(j) SE(j) + A_E AP_e(j) E(j) + A_k AP_K(j) K(j), \tag{B.3}$$

where $AP_{se}(j)$, $AP_e(j)$, and $AP_k(j)$, are the average productivity of unincorporated SE, wage employment, and computer capital for task j, respectively. In addition, A is an efficiency-augmenting technology such that A_{SE}, A_E and A_K represent the sets of skill-biased technology mapping to unincorporated SE, wage employees, and computer capital, respectively. In equation B.3, two critical assumptions are made. First, we do not distinguish between high- and low-skilled wage employment, allowing equation B.3 to consider computer capital as either a substitute for wage workers due to its linear structure or complementary, depending on the efficiency-augmenting technology A_E. This assumption is grounded by research findings from Kozeniauskas (2022), Krusell et al. (2000), and Autor et al. (2003), which suggest that advancements in capital technology enable capital to act as a substitute for low-skilled labor while complementing the high-skilled labor force.

Second, although the investment in computer capital may displace tasks traditionally performed by unincorporated SE, we posit that computerization can also bring benefits to this sector. The decreasing price of computer capital (Salgado; 2020) enables unincorporated SE to harness computerization to improve their operations. This is evident in the adoption of online reservation systems

[38] We do not specify the task types into manual, routine, and abstract as Acemoglu and Autor (2011) do. This setting simplifies the analysis for our purpose. Increasing the task types from two to three (or more) does not alter our conclusions under this influence mechanism. However, increasing the types of tasks in the analysis greatly complicates the comparative statistics given the complexity of inverting the N by N matrix ($N \geq 3$) for three or more than three threshold values. The analysis distinguishing three tasks (manual, routine, and abstract) is available upon request.

or delivery services by restaurants, as well as the expansion of small independent video game studios. However, as suggested by Davis and Haltiwanger (2014), computerization tends to provide a productivity advantage to larger firms within the economy, attributed to their scale and better access to financing. With a linear production function, each task can be undertaken by any factor of production based on the comparative advantage of different skills across tasks. We argue that computerization introduces more efficiency-augmenting technology into larger-scale entities, thereby benefiting wage workers working in such firms to a greater extent. Given that unincorporated SE typically operates on a smaller scale, we assume that, on average, $A_E \geq A_{SE}$.

Similar to the rank order across individuals in the supply context, here we rank the comparative advantage of skill groups across different tasks. Tasks characterized by lower levels of routine and standardization, which are better suited for self-employment, are positioned closer to 0. Conversely, tasks that follow well-defined procedures and are repetitive in nature, which are more suitable for computer capital, are closer to the opposite end of the spectrum. While we do not explicitly categorize wage workers into high- and low-skilled groups, we implicitly recognize different types of wage workers distributed across the middle of the spectrum. Manual/abstract-type wage workers are positioned closer to 0, while routine-type workers are closer to 1. Specifically, AP_{se}/AP_e and AP_e/MP_k strictly decrease in j such that $lim_{j\to 0}AP_{se}/AP_e = \infty$ and $lim_{j\to 1}AP_e/AP_k = 0$. This assumption implies that at the two extremes, there are always some tasks taken by the unincorporated SE given the non-routine nature of those tasks, and some tasks, given their highly automated and routinized nature, are exclusively performed by computer capital. In addition, there exist two nonarbitrage conditions that allocate the tasks to different skill groups and production factors. We relegate the discussions of two nonarbitrage conditions in Appendix B.3.1. The existence of two nonarbitrage conditions implies:

$$(\frac{E}{SE})_{demand} = (\frac{j^k - j^{se}}{j^{se}})(\frac{A_{SE}AP_{se}}{A_E AP_e}) = (\frac{j^k - j^{se}}{j^{se}})(\frac{w}{\pi})^{-1}, \tag{B.4}$$

and

$$(\frac{K}{E})_{demand} = (\frac{1 - j^k}{j^k - j^{se}})(\frac{A_E AP_e}{A_K AP_k}) = (\frac{1 - j^k}{j^k - j^{se}})(\frac{r}{w})^{-1}, \tag{B.5}$$

where $w = p(j)A_E AP_e$ and $r = p(j)A_K AP_k$ are the average returns for wage workers and computer capital and $\pi = p(j)AP_{se}$ is the average return for unincorporated SE. As shown in Equations (B.4) and (B.5), both relative demand curves E/SE and K/E are *downward sloping* with increasing payoff ratios w/π and r/w. Moreover, the relative values of thresholds j^{se} and j^k determine the shifts of the two relative demand curves and play key roles in the comparative analysis.

B.2.2 Market Equilibrium and Comparative Statistics

The market clearing condition can be derived from the relative skill-supply curve in Equation (B.1) and demand curves in Equations (B.4) and (B.5). Since individuals choose their skill supply endogenously based on the payoff ratio (w/π), as shown in Equation (B.1), we have

$$E = \int_0^{i^\star} e_i di = \Gamma_E(\ln \frac{w}{\pi}) = \Gamma_E(\ln(A_E/A_{SE}) - \beta_{se}(j^{se})) \tag{B.6}$$

and

$$SE = \int_{i^\star}^1 se_i di = \Gamma_{SE}(\ln \frac{w}{\pi}) = \Gamma_{SE}(\ln(A_E/A_{SE}) - \beta_{se}(j^{se})), \tag{B.7}$$

where $\beta_{se}(j^{se}) = \ln AP_{se} - \ln AP_e$ is the average productivity premium between individuals in the unincorporated SE and wage workers, which is strictly decreasing with an increasing task level j^{se}, denoted as $\beta'_{se}(j) < 0$. Equations (B.6) and (B.7) present the market-clearing conditions that allow us to analyze the comparative statistics with computerization. We present the mathematical details of the analytical solution and minor comparative statistics outcomes in Appendix B.3.2, and report the effects of computerization on the unincorporated SE share through two opposite forces: the restructuring effect and the efficiency-augmenting effect.

The *restructuring effect* is derived from the expansion of computer capital (K), which can negatively affect the E/SE ratio as shown in Equation (B.8).

$$\frac{d\ln(\frac{E}{SE})}{d\ln K} = \frac{\frac{\beta'_{se}W}{(j^k - j^{se})}}{|\Delta|} < 0, \tag{B.8}$$

where $\beta'_{se} < 0$, $W > 0$, and Δ is a positive definite matrix defined in Appendix B.3.2.[39] Equation (B.8) shows that computerization (an increase in computer capital) leads to a decrease in the E/SE ratio. It also indicates that while there is a reduction in the tasks carried out by wage workers, they are not completely displaced by computer capital and self-employment ($j^k - j^{se} \neq 0$). This suggests that computer capital substitutes tasks that were previously performed by wage workers. Some of the displaced workers transition into unincorporated SE, while others continue to seek wage employment opportunities within smaller-scale business entities, such as those operating as unincorporated SE businesses.

Computerization not only increases the usage of computer capital but also has a potential *efficiency-augmenting effect*. We contend that while both smaller-scale and larger-scale businesses can benefit from efficiency augmentation (as indicated in Equation B.3), larger-scale firms stand to gain more than smaller businesses Davis and Haltiwanger (2014). Given that unincorporated SE tends to be, on average, smaller in scale, we assume $A_E \geq A_{SE}$, or $\ln(A_E/A_{SE}) > 0$. Similar to the previous argument, the responses of the E/SE thresholds given the change in A_E/A_{SE} are

$$\frac{d\ln(\frac{E}{SE})}{d\ln(A_E/A_{SE})} = \frac{\frac{1}{j^k - j^{se}}\Lambda_2 - \frac{1}{j^{se}}\Lambda_1}{|\Delta|} > 0, \tag{B.9}$$

where $\Lambda_1 < 0$ and $\Lambda_2 > 0$.[40] Equation (B.9) shows that the elasticity of the relative efficiency-

[39] $W = 1 - \frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)} + \frac{\Gamma'_E(.)}{\Gamma_E(.)}$, where $\Gamma'_E(.) > 0$ and $\Gamma'_{SE}(.) < 0$.

[40] $\Lambda_1 = \beta'_k W - \frac{1}{1-j^k}W + \frac{1}{j^k - j^{se}}\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)} < 0$ and $\Lambda_2 = -\beta'_{se}W - \beta'_k W + \frac{1}{j^{se}}(1 + \frac{\Gamma'_E(.)}{\Gamma_E(.)}) + \frac{1}{1-j^k}W > 0$, where $\beta_k = \ln AP_e - \ln AP_k$ is the average productivity premium between wage workers and computer capital. This premium is

augmenting effect on the relative E/SE is positive.

The two opposite forces of the restructuring and efficiency-augmenting effects together determine the impact of computerization on the change in the unincorporated SE share. Specifically, the restructuring effect from computerization decreases the E/SE ratio, which is equivalent to an increase in the unincorporated SE share. On the other hand, the efficiency-augmenting effect increases the E/SE ratio or, put another way, decreases the unincorporated SE share.

B.2.3 Net Effect

Assume that the computerization enhances the relative efficiency through $A_E/A_{SE} = K^{\eta}$, where η represents the elasticity of the relative efficiency-augmenting from computerization. The elasticity measures the percentage of relative efficiency improvement in wage workers compared to unincorporated SE individuals resulting from a one percent increase in computer capital ($K > 1$). When $0 \leq \eta \leq 1$, we observe that $A_E \geq A_{SE}$, and when $\eta < 0$, we observe that $A_E < A_{SE}$. Given our assumption that wage workers would experience greater efficiency gains from computerization compared to unincorporated self-employed individuals, $\eta \geq 0$ is deemed more probable.

For instance, in industries when the elasticity of relative efficiency-augmenting (η) equals 1, wage workers' productivity gains from computer capital are fully magnified compared to unincorporated SE individuals, resulting in a ratio of $A_E/A_{SE} = K$. Conversely, in industries where $\eta = 0$, the productivity boost from computers for wage workers matches that of unincorporated SE individuals, resulting in a ratio of $A_E/A_{SE} = 1$. In this context, we classify industries with incorporated firms as absorbing more productivity improvements compared to unincorporated SE when $\eta \to 1$. This suggests significant efficiency gains for wage workers on large incorporated firms in contrast to unincorporated SE individuals resulting from the adoption of computer capital. Conversely, an industry tends towards similar productivity improvements through computerization for both wage workers and unincorporated SE individuals when $\eta \to 0$, indicating minimal productivity enhancements for wage workers relative to unincorporated SE individuals from adopting computers.

Given $A_E/A_{SE} = K^{\eta}$, the efficiency-augmenting effect (Equation B.9) can be expressed as $\frac{d\ln(\frac{E}{SE})}{d\ln(\frac{A_E}{A_{SE}})} = \frac{1}{\eta}\frac{d\ln(\frac{E}{SE})}{d\ln K}$. We can rewrite it as

$$\frac{d\ln(\frac{E}{SE})}{d\ln K} = \eta\frac{d\ln(\frac{E}{SE})}{d\ln(\frac{A_E}{A_{SE}})}. \tag{B.10}$$

The net effect of the efficiency-augmenting effect (Equation B.9) and the restructuring effect (Equation B.8), resulting from a one percent change in K, can be described as follows:

strictly decreasing with j, denoted as $\beta_k' < 0$. For more detailed notations, please refer to Appendix B.3.2.

$$
\begin{aligned}
&\underbrace{\frac{d\ln(\frac{E}{SE})}{d\ln K}}_{\text{efficiency-augmenting effect}} + \underbrace{\frac{d\ln(\frac{E}{SE})}{d\ln K}}_{\text{restructuring effect}} \\
= &\frac{\frac{1}{j^k-j^{se}}(\eta\Lambda_2+\beta'_{se}W)-\frac{1}{j^{se}}(\eta\Lambda_1)}{|\Delta|} \\
= &\frac{\frac{1}{j^k-j^{se}}(1-\eta)\beta'_{se}W+\frac{1}{j^k-j^{se}}(\eta\Lambda_3)-\frac{1}{j^{se}}(\eta\Lambda_1)}{|\Delta|} \gtreqless 0,
\end{aligned} \tag{B.11}
$$

where $\Lambda_1 < 0$, $\Lambda_3 > 0$, and $\beta'_{se} < 0$.[41] The net effect of these two effects is ambiguous.

Equation (B.11) demonstrates that in industries where workers' relative productivity is fully enhanced by computer capital ($\eta = 1$), the net effect is positive, given by:

$$
\begin{aligned}
&\underbrace{\frac{d\ln(\frac{E}{SE})}{d\ln K}}_{\text{efficiency-augmenting effect}} + \underbrace{\frac{d\ln(\frac{E}{SE})}{d\ln K}}_{\text{restructuring effect}} \\
= &\frac{\frac{1}{j^k-j^{se}}(\Lambda_3)-\frac{1}{j^{se}}(\Lambda_1)}{|\Delta|} > 0.
\end{aligned}
$$

This indicates that the efficiency-augmenting effect dominates the restructuring effect resulting from computerization, leading to a decrease in the unincorporated SE share. Essentially, since the productivity of wage workers in larger incorporated firms experiences a greater enhancement with the adoption of computerization, more individuals are inclined to pursue employment within such firms rather than opting for unincorporated SE.

However, as the relative elasticity of efficiency-augmenting (η) decreases, the relative efficiency-augmenting effects on wage workers diminish and become less dominant compared to the restructuring effects. In industries where computerization does not significantly complement wage workers' productivity compared to unincorporated SE individuals' (small η), as shown in Equation (B.11), the restructuring effects can outweigh the efficiency-augmenting effect. Finally, in industries where the efficiency-augmenting effects of computerization are the same for both wage workers in incorporated firms and unincorporated SE individuals ($\eta = 0$), the overall effect turns negative, resulting in a rise in the proportion of unincorporated SE individuals, illustrated as follows:

$$
\begin{aligned}
&\underbrace{\frac{d\ln(\frac{E}{SE})}{d\ln K}}_{\text{efficiency-augmenting effect}} + \underbrace{\frac{d\ln(\frac{E}{SE})}{d\ln K}}_{\text{restructuring effect}} \\
= &\frac{\frac{1}{j^k-j^{se}}\beta'_{se}W}{|\Delta|} < 0.
\end{aligned} \tag{B.12}
$$

[41] $\Lambda_3 = -\beta'_k W + \frac{1}{j^{se}}(1+\frac{\Gamma'_E(.)}{\Gamma_E(.)}) + \frac{1}{1-j^k}W > 0$, given $\beta'_k < 0$, $\Gamma'_E > 0$, and $W > 0$.

Intuitively, when productivity enhancements are equal for both wage workers and unincorporated SE individuals, the decision-making process is no longer solely determined by the returns from productivity. Individuals become indifferent between choosing to be a wage worker or engaging in unincorporated SE. In such instances, the restructuring effect will take precedence in the relationship between computerization and the change in unincorporated SE.

B.3 A Simple Model: Detailed Derivation

B.3.1 Two non-arbitrage conditions

The first nonarbitrage condition involves the case where, for task j^{se}, the profitability is identical when it is performed by either unincorporated SE or wage workers working in any business entities (including those working in unincorporated SE). Specifically, for task j^{se}, the nonarbitrage condition is

$$p(j^{se})A_{SE}AP_{se}(j^{se})se(j^{se}) = p(j^{se})A_E AP_e e(j^{se}), \tag{B.1}$$

where $p(j)$ is the price of task j. Any task $j < j^{se}$ is produced by the unincorporated SE since the wage workers do not have a comparative advantage in performing any tasks $j < j^{se}$. On the other hand, any task $j > j^{se}$ is produced by wage workers or computer capital. With a similar argument, the second nonarbitrage condition is

$$p(j^k)A_E MP_e(j^k)e(j^k) = p(j^k)A_K MP_k k(j^k). \tag{B.2}$$

Any task $j > j^k$ is produced by computer capital since those tasks are more profitable when automated, and task $j < j^k$ is performed by either wage workers or unincorporated self-employment. With the two thresholds j^{se} and j^k from the no-arbitrage conditions, we have the allocation of three skill groups/production factors such that

$$\begin{aligned}
SE &= \int_0^{j^{se}} se(j)dj = se(j)(j^{se}) \\
E &= \int_{j^{se}}^{j^k} e(j)dj = e(j)(j^k - j^{se}) \\
K &= \int_{j^k}^{1} k(j)dj = k(j)(1 - j^k).
\end{aligned}$$

The above equations hold if the factors are equally distributed across tasks performed by the same type of skill group/production factor such that $se(j) = se(j')$, $e(j) = e(j')$, and $k(j) = k(j')$, where tasks j and j' are performed by the same type of factor. Acemoglu and Autor (2011) show that the condition holds because of the unitary elasticity of technical substitutions between tasks under Cobb-Douglas technology. Replacing them into the non-arbitrary conditions in Equations

(B.1) and (B.2), we have

$$\frac{A_{SE}AP_{se}SE}{j^{se}} - \frac{A_E AP_e E}{j^k - j^{se}} = 0, \tag{B.3}$$

and

$$\frac{A_E AP_e E}{j^k - j^{se}} - \frac{A_K AP_k K}{1 - j^k} = 0. \tag{B.4}$$

Equations (B.3) and (B.4) can be rewritten as

$$\frac{E}{SE} = (\frac{j^k - j^{se}}{j^{se}})(\frac{A_{SE}AP_{se}}{A_E AP_e}) = (\frac{j^k - j^{se}}{j^{se}})(\frac{w}{\pi})^{-1}, \tag{B.5}$$

and

$$\frac{K}{E} = (\frac{1 - j^k}{j^k - j^{se}})(\frac{A_E AP_e}{A_K AP_k}) = (\frac{1 - j^k}{j^k - j^{se}})(\frac{r}{w})^{-1}, \tag{B.6}$$

where $w = p(j)A_E AP_e$ and $r = p(j)A_K AP_k$ are the average returns for wage workers and computer capital and $\pi = p(j)AP_{se}$ is the average return for unincorporated SE.

B.3.2 Market clearing condition and comparative statistics

We present the mathematical details of the market clearing condition. The market clearing condition can be derived from the relative skill-supply curve in Equation (B.1) and demand curves in Equations (B.4) and (B.5). Since individuals choose their skill supply endogenously based on the payoff ratio (w/π), as shown in Equation (B.1), we have

$$E = \int_0^{i^\star} e_i di = \Gamma_E(\ln \frac{w}{\pi}) = \Gamma_E(\ln(A_E/A_{SE}) - \beta_{se}(j^{se}))$$

and

$$SE = \int_{i^\star}^{1} se_i di == \Gamma_{SE}(\ln \frac{w}{\pi}) = \Gamma_{SE}(\ln(A_E/A_{SE}) - \beta_{se}(j^{se})).$$

The above equations are derived from the fact that $w/\pi = (A_E AP_e)/(A_{SE}AP_{se})$, where $\beta_{se}(j^{se}) = \ln AP_{se} - \ln AP_e$ is the average productivity premium between individuals in the unincorporated SE and wage employees, which is strictly decreasing with an increasing task level j^{se}. Given this setting, we have $\Gamma'_E(.) > 0$ and $\Gamma'_{SE}(.) < 0$. By taking logs on the two relative demand curves ((B.4) and (B.5)), we have

$$\ln \Gamma_{SE}(\ln(A_E/A_{SE}) - \beta_{se}(j^{se})) - \ln \Gamma_E(\ln(A_E/A_{SE}) - \beta_{se}(j^{se})) + \ln(j^k - j^{se}) - \ln(j^{se}) - \ln(A_E/A_{SE}) + \beta_{se}(j^{se}) = 0, \tag{B.7}$$

and

$$\ln \Gamma_E(\ln(A_E/A_{SE}) - \beta_{se}(j^{se})) - \ln K + ln(1 - j^k) - \ln(j^k - j^{se}) + \ln(A_E/A_k) + \beta_k(j^k) = 0. \tag{B.8}$$

Similarly, we define $\beta_k(j^k) = \ln AP_e - \ln AP_k$, which is strictly decreasing with j^k.

Equations (B.7) and (B.8) present the market clearing conditions that allow us to analyze the comparative statistics with computerization. As mentioned in the introduction, computerization has both restructuring and efficiency-augmenting effects on the ratio of wage workers to unincorporated SE (E/SE). The restructuring effect occurs when increasing computer capital leads to the substitution of wage workers. To show the restructuring effect of computerization, we consider the responses of E/SE to a change in K by totally differentiating Equations (B.7) and (B.8):

$$\begin{bmatrix} \beta'_{se}(1-\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)}+\frac{\Gamma'_{E}(.)}{\Gamma_{E}(.)})-\frac{1}{j^k-j^{se}}-\frac{1}{j^{se}} & \frac{1}{j^k-j^{se}} \\ -\frac{\Gamma'_{E}(.)}{\Gamma_{E}(.)}\beta'_{se}+\frac{1}{j^k-j^{se}} & \beta'_k-\frac{1}{1-j^k}-\frac{1}{j^k-j^{se}} \end{bmatrix} \begin{bmatrix} dj^{se} \\ dj^k \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} d\ln K. \tag{B.9}$$

Given the properties that $\beta'_{se} < 0$, $\beta'_k < 0$, $\Gamma'_E(.) > 0$, and $\Gamma'_{SE}(.) < 0$, the determinant of the matrix on the left-hand side is positive definite and denoted by Δ. Therefore, the comparative statistics with an increase in computer capital (K) are

$$\frac{dj^{se}}{d\ln K} = \frac{\frac{-1}{j^k-j^{se}}}{|\Delta|} < 0,$$

$$\frac{dj^k}{d\ln K} = \frac{\beta'_{se}W-\frac{1}{j^k-j^{se}}-\frac{1}{j^{se}}}{|\Delta|} < 0,$$

and

$$\frac{d(j^k-j^{se})}{d\ln K} = \frac{\beta'_{se}W-\frac{1}{j^{se}}}{|\Delta|} < 0,$$

where $W = 1-\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)}+\frac{\Gamma'_{E}(.)}{\Gamma_{E}(.)} > 0$. Given the responses of thresholds with the change in computer capital and Equation (B.4), we can therefore derive the elasticity of E/SE in the context of computerization as

$$\frac{d\ln(\frac{E}{SE})}{d\ln K} = \frac{d\ln(j^k-j^{se})}{d\ln K} - \frac{d\ln(j^{se})}{d\ln K} = \frac{\frac{\beta'_{se}W}{(j^k-j^{se})}}{|\Delta|} < 0. \tag{B.10}$$

Equation (B.10) shows that computerization (an increase in computer capital) leads to a decrease in the E/SE ratio.

Computerization not only increases the usage of computer capital but also has a potential efficiency-augmenting effect. Specifically, incorporated firms utilize computerization to facilitate the productivity of wage workers. This efficiency-augmenting effect reflects the A_E/A_{SE}. Note that we assume that there is no efficiency-augmenting term for individuals in the unincorporated SE given that it is less likely that they have a level of computerization comparable to that in incorporated companies. Similar to the previous argument, the responses of the thresholds given the change in A_E/A_{SE} are

$$\begin{bmatrix} \beta'_{se}(1-\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)}+\frac{\Gamma'_{E}(.)}{\Gamma_{E}(.)})-\frac{1}{j^k-j^{se}}-\frac{1}{j^{se}} & \frac{1}{j^k-j^{se}} \\ -\frac{\Gamma'_{E}(.)}{\Gamma_{E}(.)}\beta'_{se}+\frac{1}{j^k-j^{se}} & \beta'_k-\frac{1}{1-j^k}-\frac{1}{j^k-j^{se}} \end{bmatrix} \begin{bmatrix} dj^{se} \\ dj^k \end{bmatrix} = \begin{bmatrix} W \\ -(1+\frac{\Gamma'_{E}(.)}{\Gamma_{E}(.)}) \end{bmatrix} d\ln A_E. \tag{B.11}$$

Given the properties that $\beta'_{se} < 0$, $\beta'_k < 0$, $\Gamma'_E(.) > 0$, and $\Gamma'_{SE}(.) < 0$, the comparative statistics with an increase in relative efficiency-augmenting (A_E/A_{SE}) are

$$\frac{dj^{se}}{d\ln(A_E/A_{SE})} = \frac{\beta'_k W-\frac{1}{1-j^k}W+\frac{1}{j^k-j^{se}}\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)}}{|\Delta|} = \frac{\Lambda_1}{|\Delta|} < 0,$$

$$\frac{dj^k}{d\ln(A_E/A_{SE})} = \frac{-\beta'_{se}W + \frac{1}{j^{se}}(1+\frac{\Gamma'_E(.)}{\Gamma_E(.)}) + \frac{1}{j^k - j^{se}}\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)}}{|\Delta|} \gtreqless 0,$$

and

$$\frac{d(j^k - j^{se})}{d\ln(A_E/A_{SE})} = \frac{-\beta'_{se}W - \beta'_k W + \frac{1}{j^{se}}(1+\frac{\Gamma'_E(.)}{\Gamma_E(.)}) + \frac{1}{1-j^k}W}{|\Delta|} = \frac{\Lambda_2}{|\Delta|} > 0.$$

Note that since we allow endogenous choice of both skill supply and demand, the threshold change in j^k is less clear than in the inelastic labor supply environment. This is because with the increase in efficiency-augmenting technology, some originally self-employed individuals shift their profession to join incorporated companies (appearing in the negative term $\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)}\frac{1}{j^k - j^{se}}$ in the $\frac{dj^k}{d\ln A_E}$ equation), further lowering j^{se}, which might potentially lower j^k indirectly. Nevertheless, the efficiency-augmenting hypothesis holds since

$$\frac{d\ln(\frac{E}{SE})}{d\ln(A_E/A_{SE})} = \frac{\frac{1}{j^k - j^{se}}\Lambda_2 - \frac{1}{j^{se}}\Lambda_1}{|\Delta|} > 0, \tag{B.12}$$

where $\Lambda_1 = \beta'_k W - \frac{1}{1-j^k}W + \frac{1}{j^k - j^{se}}\frac{\Gamma'_{SE}(.)}{\Gamma_{SE}(.)} < 0$ and $\Lambda_2 = -\beta'_{se}W - \beta'_k W + \frac{1}{j^{se}}(1+\frac{\Gamma'_E(.)}{\Gamma_E(.)}) + \frac{1}{1-j^k}W > 0$. Equation (B.12) shows that the elasticity of the efficiency-augmenting effect on the relative E/SE is positive. This hypothesis assumes that computerization complements the skills of wage employees working in incorporated firms. With the efficiency-augmenting advantages of computerization, incorporated firms can now conduct more tasks originally performed by unincorporated SE individuals. In addition, some originally unincorporated SE individuals become wage employees given the increase in the opportunity cost of not working in incorporated firms.

The two opposite forces of the restructuring and efficiency-augmenting effects together determine the impact of computerization on the change in the unincorporated SE share. Specifically, the restructuring effect from computerization decreases the E/SE ratio, which is equivalent to an increase in the unincorporated SE share. On the other hand, the efficiency-augmenting effect increases the E/SE ratio or, put another way, decreases the unincorporated SE share.

C Supplementary Figures and Tables to Section 4

C.1 Map EWCS Industry Category to US Census Data

The EWCS's industry and occupation classifications are different from the US Census data. Because there is no available crosswalk between these two surveys, we manually map the industry and occupation classifications of the EWCS to the US 1950 Census survey.

Table C.1: The occupation mapping of 1990, 2000 EWCS to 1950 IPUMS

Panel A. Occupation mapping of 1990 EWCS to 1950 IPUMS			
1990 EWCS Classification	Survey Code	IPUMS Classification	Survey Code
Professional ((lawyer, medical practitioner, accountant, etc)	3	Professional	0-99
Farmer	1	Farmers	100-123
		Farm laborers	810-840
Fisherman	2	Farmers	100-123
		Farm laborers	810-840
Owners of shops or companies, craftsmen	4	Craftsmen	500-595
Employed professional ((lawyer, medical practitioner, accountant, etc)	5	Professional	0-99
General management	6	Managers, officials, and proprietors	200-290
Middle management	7	Managers, officials, and proprietors	200-290
Other office employees	8	Clerical	300-390
Non-office employees, non manual worker	9	Sales	400-490
Supervisors	10	Managers, officials, and proprietors	200-290
Skilled manual worker	11	Operatives	600-690
Other manual worker	12	Service	700-790
		Laborers	910-970
Panel B. Occupation mapping of 2000 EWCS to 1950 IPUMS			
2000 EWCS Classification	Survey Code	IPUMS Classification	Survey Code
Legislators, senior officials and managers	1	Managers, officials, and proprietors	200-290
Professionals	2	Professional	0-99
Technicians and associate professionals	3	Professional	0-99
Clerks	4	Clerical	300-390
Service workers and shop and market sales workers	5	Service	700-790
Skilled agricultural and fishery workers	6	Farmers	100-123
		Farm laborers	810-840
Craft and related trades worker	7	Craftsmen	500-595
Plant and machine operators and assemblers	8	Operatives	600-690
		Laborers	910-970
Elementary occupations	9	Sales	400-490

To extract the industry and occupation information, we rely on the following survey questions from the 1990 EWCS: "What is the main business activity of the establishment (factory, office...) where you work? (a1r3)" and "What is your occupation? (d17r)"; and, in 2000 EWCS: "What is the main business activity of the establishment (factory, office...) where you work? (q5r)" and "1st-level ISCO codes (isco)". These two questions in each survey year provide roughly the one-digit industry and occupation classification. We present the crosswalk we used to match the 1990, 2000 EWCS code to the 1950 IPUMS Census occupation code in Table C.1 (occupation mapping), and in Table C.2 (3-digit IPUMS industry code to the 1990 and 2000 EWCS industry codes).

Table C.2: The industry mapping of 1990, 2000 EWCS to 1950 IPUMS

Panel A. Industry mapping of 1990 EWCS to 1950 IPUMS	
1991 EWCS Classification	IPUMS Industry Code
Agriculture and fishing	10-32
Energy/water/ extraction and processing/ chemistry	40-50, 180-222
Metal manufacture	270-370
Other manufacture industries	100-172, 230-260, 371-392
Building and civil engineering	60
	882
Distributive trades, hotels, catering, repairs	500-691
Transport and communication	400-472
Banking and finance, insurance, business services,	700-712
Other services (including public administration)	721-881, 890-932
Panel B. Industry mapping of 2000 EWCS to 1950 IPUMS	
2000 EWCS Classification	IPUMS Industry Code
Agriculture and fishing	10-32
Manufacturing and Mining	40-332, 351-392
Electricity, gas and water supply	340-350
Construction	60
Wholesale and retail trade	500-691
Hotels and restaurants	761-791
Transport and communications	400-472
Financial intermediation	700-711
Real estate	712
Public administration and defence	900-932
Education and health	721-760, 800-893

D Supplementary Figures and Equations to Section 5

D.1 Routine-intensive and Abstract-intensive Occupation Measures

Autor and Dorn (2013) propose a framework in which we assume that workers perform routine, abstract, or manual tasks, and we combine these tasks measures to create a summary measure of routine task intensity (RTI) and abstract task intensity (ATI) by occupation. These two measures are calculated as

$$\text{Routine}: \quad RTI_k = ln(T^R_{k,t}) - ln(T^M_{k,t}) - ln(T^A_{k,t})$$

$$\text{Abstract}: \quad ATI_k = ln(T^A_{k,t}) - ln(T^R_{k,t}) - ln(T^M_{k,t})$$

where T^R_k, T^M_k and T^A_k are the routine, manual, and abstract task inputs, respectively, in each occupation k at time t. By construction, RTI_k and ATI_k focus on the importance of routine and abstract tasks, respectively, in each occupation.

Following Autor and Dorn (2013), we define highly routine and abstract task-intensive occupations by identifying two sets of occupations in the top employment-weighted third of RTI_k and ATI_k, respectively, at time t. Thus, using the measures of RTI_k and ATI_k, for each CZ c, we calculate the routine task-intensive employment share (RSH_{ct}) and abstract task-intensive employment share (ASH_{ct}) as

$$RSH_{ct} = \left(\sum_{k=1}^{K} L_{ckt} \times 1\Big(RTI_k > RTI_{P66}\Big)\right)\left(\sum_{k=1}^{K} L_{ckt}\right)^{-1} \tag{D.1}$$

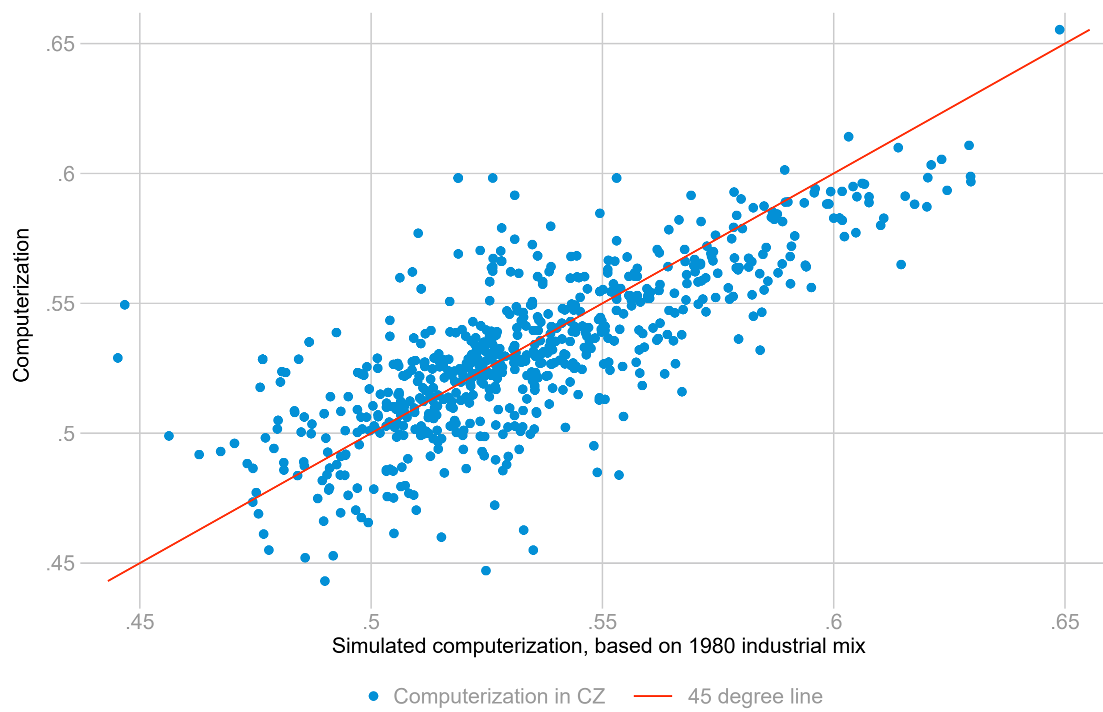
$$ASH_{ct} = \left(\sum_{k=1}^{K} L_{ckt} \times 1\Big(ATI_k > ATI_{P66}\Big)\right)\left(\sum_{k=1}^{K} L_{ckt}\right)^{-1}, \tag{D.2}$$

where L_{ckt} is employment in occupation k in CZ c at time t and the indicator function $1[\cdot]$ takes the value 1 if $QTI_k > QTI_{P66}$ and $QTI = \{RTI, ATI\}$.

D.2 Simulating Computerization

Figure D.1 depicts the survey-based data and simulated computerization levels for 2000. As shown in the figure, the survey-based data and simulated computerization levels are highly correlated and do not cluster on either side of the 45-degree line. If we look at more details, the simulated computerization growth is slightly less than the growth of the survey-based data computerization.

Figure D.1: Simulated and Survey-based Data Computerization Comparison at CZ Level



Specifically, the survey-based data population-weighted average computerization rose from 40.376 percent in 1990 to 57.372 percent in 2000, while simulated computerization increased from 40.487 percent in 1990 to 56.180 percent in 2000. This implies that the compositional change is slanted towards industries with high computerization; however, the induced CZ-level computerization is negligible.

D.3 Heterogeneous Effect of Computerization on Wage

We adopt the quantile IV method proposed by Chernozhukov and Hansen (2008) to estimate the relationship between computerization and top-quantile wages. Specifically, we estimate computerization in a CZ on the logged real hourly wage growth of wage workers in various quantiles of the wage distribution function. To apply the instrumental variable quantile regression (IVQR) method, we estimate the equation as follows:

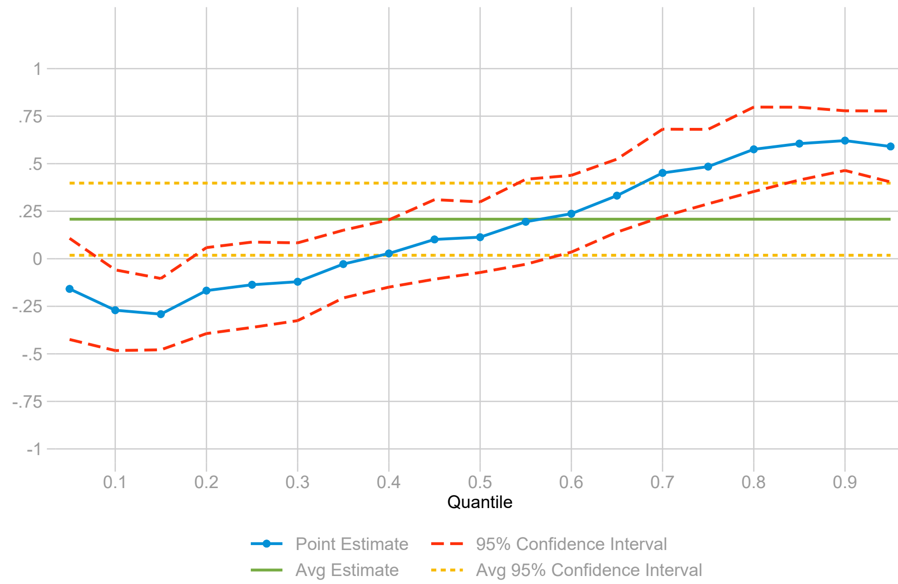
$$\Delta ln\bar{w}_g = \tilde{\gamma}_g + \tilde{\lambda}_\tau Computerization_{g_{t_0}} + \tilde{\delta} + \tilde{\phi}_s + \tilde{e}_g. \tag{D.3}$$

Note that we have changed the notation in Equation (D.3) in comparison with that in Equation (2). In this setup, group g is a given CZ in a given decade.

The variable of interest $Computerization_{g_{t_0}}$ represents group g's computerization in t_0. The dependent variable $\Delta ln\bar{w}_g$ denotes the change in log hourly wage from 1990 to 2010 for group g. The coefficient of interest $\tilde{\lambda}_\tau$ represents the impact of computerization in t_0 on wage growth during the 1990-2010 period at the τth quantile of the wage distribution. We apply the instrumental variables $\widetilde{RSH}_{g1950}$ and $\widetilde{ASH}_{g1950}$ that are used in Autor and Dorn (2013) to address the endogeneity issue.

The IVQR results are plotted in Figure D.2, which suggests that computerization substantially increases the top quantile's hourly wage by 75 percent, while the average increase is approximately 25 percent. As in Panel B of Table 7, computerization has a highly heterogeneous effect on the hourly wage. The bottom quantiles' hourly wage is decreasing, and the hourly wage increases monotonically along with computerization.

Figure D.2: Computerization and the Real Hourly Wage Growth by Percentile



References

Acemoglu, D. (2002). Technical change, inequality, and the labor market, *Journal of Economic Literature* **40**(1): 7–72.

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings, *Handbook of Labor Economics*, Vol. 4, Elsevier, pp. 1043–1171.

Acemoglu, D., Gancia, G. and Zilibotti, F. (2015). Offshoring and directed technical change, *American Economic Journal: Macroeconomics* **7**(3): 84–122.

Acemoglu, D. and Restrepo, P. (2018). Low-skill and high-skill automation, *Journal of Human Capital* **12**(2): 204–232.

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

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets, *Journal of Political Economy* **128**(6): 2188 – 2244.

Addison, J., McKinley, B. and Cotti, C. (2009). Does the minimum wage raise employment? evidence from the U.S. retail trade sector, *Labour Economics* **16(4)**: 397–408.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J. and Li, H. (2022). A theory of falling growth and rising rents, working paper 2019-11.

Allegretto, S., Dube, A. and Reich, M. (2011). Does the minimum wage really reduce teen employment? accounting for heterogeneity and selectivity in state panel data, *Industrial Relations* **50(2)**: 205–240.

Atkinson, A. B. (2007). The distribution of earnings in OECD countries, *International Labour Review* **146**(1-2): 41–60.

Autor, D. (2013). The 'task approach' to labor markets, *Journal of Labour Market Research* **46 (3)**: 185–99.

Autor, D. and Dorn, D. (2013). The growth of low-skilled service jobs and polarization of the US labor market, *American Economic Review* **103(5)**: 1553–1597.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms, *The Quarterly Journal of Economics* **135**(2): 645–709.

Autor, D. and Duggan, M. (2003). The rise in the disability rolls and the decline in unemployment, *Quarterly Journal of Economics* **118(1)**: 157–205.

Autor, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation, *Journal of Economic Perspectives* **29**(3): 3–30.

Autor, D. H. and Dorn, D. (2010). Inequality and specialization: The growth of low-skilled service employment in the United States, *Technical report*, MIT Working Paper.

Autor, D. H., Dorn, D. and Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States, *American Economic Review* **103**(6): 2121–68.

Autor, D. H., Katz, L. F. and Kearney, M. S. (2006). The polarization of the US labor market, *American Economic Review* **96**(2): 189–194.

Autor, D. H., Katz, L. F. and Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists, *The Review of Economics and Statistics* **90**(2): 300–323.

Autor, D. H., Katz, L. F. and Krueger, A. B. (1998). Computing inequality: have computers changed the labor market?, *The Quarterly Journal of Economics* **113**(4): 1169–1213.

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

Bartik, T. (1991). Who benefits from state and local economic development policies?, *W.E. Upjohn Institute* .

Basso, G., Peri, G. and Rahman, A. S. (2020). Computerization and immigration: Theory and evidence from the United States, *Canadian Journal of Economics* **53**(4): 1457 – 1494.

Blanchard, O. and Katz, L. (1992). Regional evolutions, *Brookings Papers on Economic Activity* .

Blinder, A. S. and Krueger, A. B. (2013). Alternative measures of offshorability: a survey approach, *Journal of Labor Economics* **31**(S1): S97–S128.

Bornstein, G. et al. (2018). Entry and profits in an aging economy: The role of consumer inertia, *Review of Economic Studies, forthcoming* .

Borri, N., Liu, Y. and Tsyvinski, A. (2022). The economics of non-fungible tokens, *Available at SSRN 4052045* .

Borusyak, K., Hull, P. and Jaravel, X. (2022). Quasi-experimental shift-share research designs, *The Review of Economic Studies* **89**(1): 181–213.

Brynjolfsson, E. and Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance, *Journal of Economic Perspectives* **14**(4): 23–48.

Brynjolfsson, E., Li, D. and Raymond, L. R. (2023). Generative ai at work, *Technical report*, National Bureau of Economic Research.

Burtch, G., Carnahan, S. and Greenwood, B. N. (2018). Can you gig it? an empirical examination of the gig economy and entrepreneurial activity, *Management Science* **64**(12): 5497–5520.

Byrd, T. A. and Davidson, N. W. (2003). Examining possible antecedents of IT impact on the supply chain and its effect on firm performance, *Information & Management* **41**(2): 243–255.

Cagetti, M. and De Nardi, M. (2006). Entrepreneurship, frictions, and wealth, *Journal of Political Economy* **114**(5): 835–870.

Card, D. and Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis, *The Quarterly Journal of Economics* **116**(2): 705–746.

Carneiro, P. and Lee, S. (2011). Trends in quality-adjusted skill premia in the United States, 1960-2000, *American Economic Review* **101**(6): 2309–49.

Chernozhukov, V. and Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach, *Journal of Econometrics* **142**(1): 379 – 398.

Cintio, M. D. (2022). Overeducation and R&D: Theoretical aspects and empirical evidence, *Journal of Human Capital* **16**(2): 183–232.

Costinot, A. and Vogel, J. (2010). Matching and inequality in the world economy, *Journal of Political Economy* **118**(4): 747–786.

Davis, S. J. and Haltiwanger, J. (2014). Labor market fluidity and economic performance, *Technical report*, National Bureau of Economic Research.

De Guinea, A. O. and Webster, J. (2013). An investigation of information systems use patterns: Technological events as triggers, the effect of time, and consequences for performance, *MIS Quarterly* pp. 1165–1188.

De Ridder, M. (2024). Market power and innovation in the intangible economy, *American Economic Review* **114**(1): 199–251.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2016a). Declining business dynamism: What we know and the way forward, *American Economic Review* **106**(5): 203–207.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2016b). Where has all the skewness gone? the decline in high-growth (young) firms in the us, *European Economic Review* **86**: 4–23.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2017). Declining dynamism, allocative efficiency, and the productivity slowdown, *American Economic Review* **107**(5): 322–326.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2020). Changing business dynamism and productivity: Shocks versus responsiveness, *American Economic Review* **110**(12): 3952–3990.

Decker, R., Haltiwanger, J., Jarmin, R. and Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism, *Journal of Economic Perspectives* **28**(3): 3–24.

Derenoncourt, E. (2022). Can you move to opportunity? evidence from the great migration, *American Economic Review* **112**(2): 369–408.

Dorn, D. et al. (2009). This job is "getting old": Measuring changes in job opportunities using occupational age structure, *American Economic Review* **99**(2): 45–51.

Ehrlich, I., Li, D. and Liu, Z. (2017). The role of entrepreneurial human capital as a driver of endogenous economic growth, *Journal of Human Capital* **11**(3): 310–351.

Engbom, N. et al. (2019). Firm and worker dynamics in an aging labor market, *Technical report*, Federal Reserve Bank of Minneapolis Minneapolis, MN.

Feenstra, R. C. and Hanson, G. H. (1999). The impact of outsourcing and high-technology capital on wages: estimates for the United States, 1979–1990, *The Quarterly Journal of Economics* **114**(3): 907–940.

Glaeser, E. L. (2009). Entrepreneurship and the city, *Entrepreneurship and Openness*, Edward Elgar Publishing.

Glaeser, E. L., Rosenthal, S. S. and Strange, W. C. (2010). Urban economics and entrepreneurship, *Journal of Urban Economics* **67**(1): 1–14.

Goldin, C. D. and Katz, L. F. (2009). *The race between education and technology*, Harvard University Press.

Goldsmith-Pinkham, P., Sorkin, I. and Swift, H. (2020). Bartik instruments: What, when, why, and how, *American Economic Review* **110**(8): 2586–2624.

Goos, M. and Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain, *The Review of Economics and Statistics* **89**(1): 118–133.

Goos, M., Manning, A. and Salomons, A. (2009). Job polarization in Europe, *American Economic Review* **99**(2): 58–63.

Goos, M., Manning, A. and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring, *American Economic Review* **104**(8): 2509–26.

Grilo, I. and Thurik, R. (2008). Determinants of entrepreneurial engagement levels in Europe and the US, *Industrial and Corporate Change* **17**(6): 1113–1145.

Grossman, G. M. and Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring, *American Economic Review* **98**(5): 1978–97.

Hopenhayn, H., Neira, J. and Singhania, R. (2022). From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share, *Econometrica* **90**(4): 1879–1914.

Hou, C.-K. (2013). Measuring the impacts of the integrating information systems on decision-making performance and organisational performance: an empirical study of the Taiwan semiconductor industry, *International Journal of Technology, Policy and Management* **13**(1): 34–66.

Hsieh, C.-T. and Rossi-Hansberg, E. (2023). The industrial revolution in services, *Journal of Political Economy Macroeconomics* **1**(1): 3–42.

Hunt, J. (2011). Which immigrants are most innovative and entrepreneurial? distinctions by entry visa, *Journal of Labor Economics* **29**(3): 417–457.

Hunt, J. and Nunn, R. (2022). Has US employment really polarized? a critical reappraisal, *Labour Economics* **75**: 102117.

Jiang, H. and Sohail, F. (2023). Skill-biased entrepreneurial decline, *Review of Economic Dynamics* **48**: 18–44.

Kahn, S., La Mattina, G. and MacGarvie, M. J. (2017). "misfits,""stars," and immigrant entrepreneurship, *Small Business Economics* **49**(3): 533–557.

Karahan, F., Pugsley, B. and Şahin, A. (2019). Demographic origins of the startup deficit, *Technical report*, National Bureau of Economic Research.

Kendall, T. D. and Tsui, K. (2011). The economics of the long tail, *The BE Journal of Economic Analysis & Policy* **11**(1).

Kozeniauskas, N. (2022). What's driving the decline in entrepreneurship?, *Banco de Portugal, Economics and Research Department* .

Krueger, A. B. (1993). How computers have changed the wage structure: evidence from microdata, 1984–1989, *The Quarterly Journal of Economics* **108**(1): 33–60.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V. and Violante, G. L. (2000). Capital-skill complementarity and inequality: A macroeconomic analysis, *Econometrica* **68**(5): 1029–1053.

Kuratko, D. F. and Audretsch, D. B. (2022). The future of entrepreneurship: the few or the many?, *Small Business Economics* **59**(1): 269–278.

Lamadon, T., Mogstad, M. and Setzler, B. (2022). Imperfect competition, compensating differentials, and rent sharing in the US labor market, *American Economic Review* **112**(1): 169–212.

Lefter, A. M., Sand, B. M. et al. (2011). *Job Polarization in the US: A Reassessment of the Evidence from the 1980s and 1990s*, School of Economics and Political Science, Department of Economics, Univ.

Levine, R. and Rubinstein, Y. (2013). Smart and illicit: Who becomes an entrepreneur and does it pay?, *Technical report*, National Bureau of Economic Research.

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

Marcolin, L., Miroudot, S. and Squicciarini, M. (2018). To be (routine) or not to be (routine), that is the question: a cross-country task-based answer, *Industrial and Corporate Change* **28**(3): 477–501.

Meer, J. and West, J. (2016). Effects of minimum wage on employment dynamics, *Jornal of Human Resources* **51(2)**: 500–522.

Michaels, G., Natraj, A. and Van Reenen, J. (2014). Has ICT polarized skill demand? evidence from eleven countries over twenty-five years, *The Review of Economics and Statistics* **96**(1): 60–77.

Mishel, L., Schmitt, J. and Shierholz, H. (2013). Assessing the job polarization explanation of growing wage inequality, *Economic Policy Institute, Working Paper* **11**: 2013.

Nordhaus, W. D. (2007). Two centuries of productivity growth in computing, *The Journal of Economic History* **67**(1): 128–159.

Page, M., Spetz, J. and Millar, J. (2005). Does the minimum wage affect welfare caseloads?, *Journal of Policy Analysis and Management* **24(2)**: 273–295.

Qin, N. and Kong, D. (2021). Human capital and entrepreneurship, *Journal of Human Capital* **15**(4): 513–553.

Quadrini, V. (2000). Entrepreneurship, saving, and social mobility, *Review of Economic Dynamics* **3**(1): 1–40.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J. and Sobek, M. (2020). IPUMS USA: Version 10.0 [dataset].

Salgado, S. (2020). Technical change and entrepreneurship, *Available at SSRN 3616568* .

Sima, D. (2023). The long-term effects of external examinations on entrepreneurship: Evidence from global entrepreneurship monitor, *Journal of Human Capital* **17**(3): 387–433.

Sobel, R. S. and Dean, A. M. (2008). Has Wal-mart buried mom and pop?: The impact of Wal-mart on self-employment and small establishments in the united states, *Economic Inquiry* **46**(4): 676–695.

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure, *Journal of Labor Economics* **24**(2): 235–270.

Stock, J. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression*, Cambridge University Press, New York, pp. 80–108.

Sun, J. (2017). The effect of information technology on IT-facilitated coordination, IT-facilitated autonomy, and decision-makings at the individual level, *Applied Economics* **49**(2): 138–155.

Tinbergen, J. (1974). Substitution of graduate by other labour, *Kyklos: International Review for Social Sciences* .

Tolbert, C. and Killian, M. (1987). Labor market areas for the United States, *Washington, DC: Economic Research Service, US Department of Agriculture. Staff Report No AGES870721* .

Tolbert, C. and Sizer, M. (1996). U.S. commuting zones and labor market areas: A 1990 update, *Washington, DC: Economic Research Service, US Department of Agriculture. Staff Paper 9614* .

Welch, F. (1973). Black-white differences in returns to schooling, *American Economic Review* **63**(5): 893–907.