Under the Hood of the Routine Share Decline*

Nir Jaimovich
University of California, San Diego and *NBER, CEPR*
nijaimovich@ucsd.edu

Miao Ben Zhang
University of Southern California
Miao.Zhang@marshall.usc.edu

Nicolas Vincent
HEC Montréal
nicolas.vincent@hec.ca

September 2023

Abstract

Using establishments' occupational data, we quantify the role of entrants, exiters, and incumbents in driving the decline in the share of routine occupations (R-share) in the U.S. First, entrants have a higher R-share than incumbents, casting doubt on a "creative destruction" mechanism whereby entrants drive this decline. Second, exiters have a higher R-share than their peers, supporting a "positive selection" mechanism. Finally, as incumbents age, they experience a fall in their R-share, which is not due to their size, consistent with the "technology adoption" mechanism. Quantitatively, we show that incumbents are the primary drivers of the aggregate decline in R-share.

Keywords: Routine occupations, Establishments' dynamics.

*Vincent gratefully acknowledges financial support from the Fondation HEC Montréal and the Social Sciences and Humanities Research Council of Canada (grant number: 435 − 2022 − 0800). Jaimovich gratefully acknowledges financial support from the Swiss National Science Foundation (grant number: 100018 − 197799).

1 Introduction

The decline in employees' share working in routine occupations (*R-share* hereafter), known as job polarization, has been at the center of recent discussions (e.g., Autor et al. (2006), Goos and Manning (2007), and Acemoglu and Autor (2011)).

Evidence on the R-share's evolution at the establishment level is limited. Using U.S. administrative micro-data of establishments' occupational employment, we address this gap. This research furthers our knowledge on the R-share decline and how establishments modify their employment.

We document a decline over time in the R-share of incumbent, entering, and exiting establishments. For incumbents, this decline over their life cycle holds even after controlling for size and across different cohorts. While new cohorts of entrants exhibit a lower R-share than prior cohorts, entrants have a higher R-share than incumbents, even several years after birth, casting doubt on a "creative destruction" channel. Exiters, on the other hand, display a higher R-share than incumbents, both upon exit as well as in the years leading to their exit. This implies a negative, albeit very small, contribution to the R-share through selection. Ultimately, our findings point to the key role of incumbents in driving the decline in the R-share.

Studies most related to ours include Heyman (2016), Bockerman et al. (2019), and Harrigan et al. (2021), all using micro-data of occupation at the establishment level. Heyman (2016) uncovers within-firm job polarization in Sweden without addressing the entry/exit margin. Bockerman et al. (2019) finds significant within-firm adjustments for the middle education group in Finland (a proxy for Routine occupations), and unlike us, sees the entry margin contributing to the R-share decline. Harrigan et al. (2021) identifies firm composition changes, and not within-firm adjustments, as the primary driver of polarization in France. In contrast, we find within-establishment adjustments in the U.S. pivotal to the R-share decline. Overall, our contribution is to document the dynamics of entry and exit in the U.S. and enrich the analysis regarding the evolution of the R-share within incumbents.

2 R-share over time and across establishments

Our administrative data tracks occupational-level employment in approximately 1.2 million U.S. establishments stratified to represent the economy from 1988 to 2013. Online Appendix A.1 provides more details.

We measure each establishment's share of routine-task labor by following the standard definition in the literature (e.g., the definition of Jaimovich and Siu (2020) described in Appendix A.2): Our main variable is an establishment's share of total employment in routine-task labor **(R-share)**:

$$\text{R-share}_{i,t} = \frac{\sum_o 1[o \in R] * emp_{o,i,t}}{\sum_o emp_{o,i,t}}. \tag{1}$$

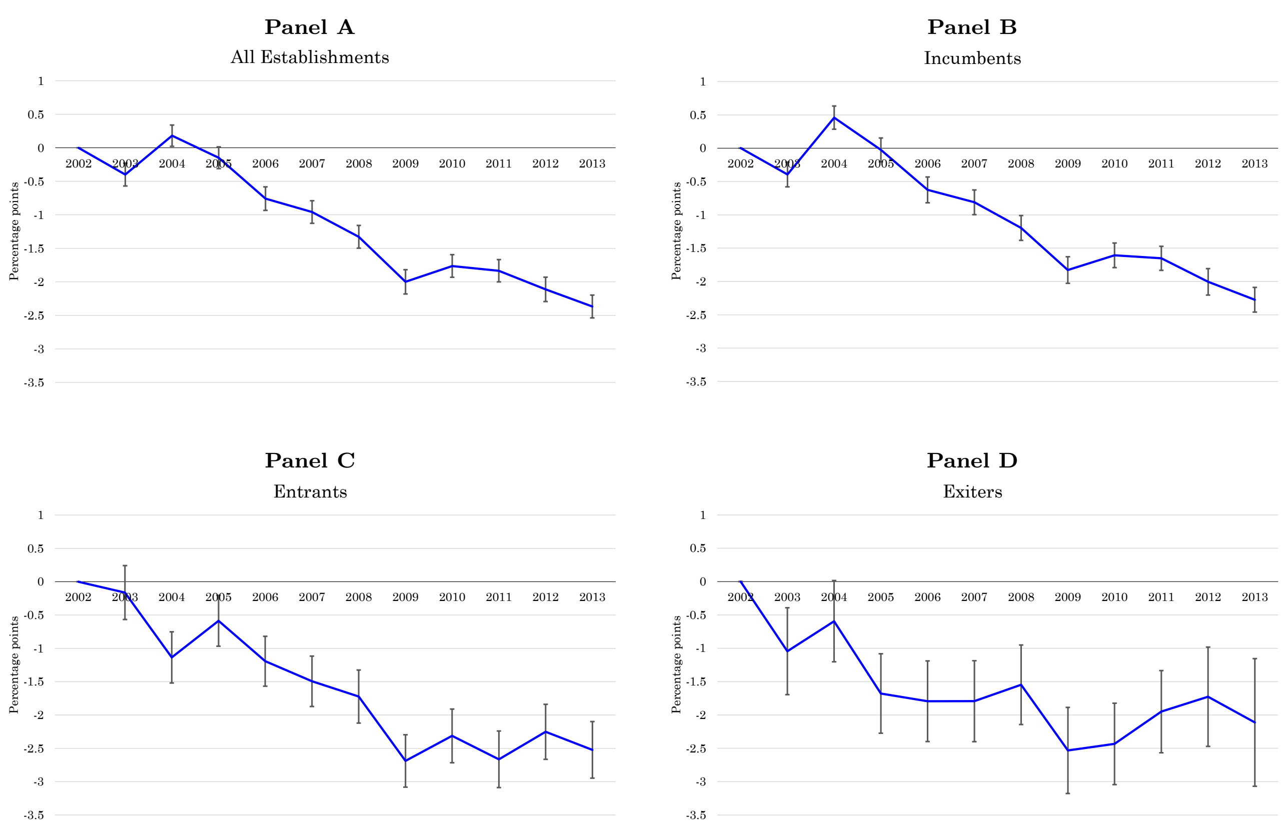
where o, i and t respectively refer to an occupation, establishment, and year.

We begin by regressing establishment-level R-share on year dummies, while controlling for NAICS3 fixed effects for establishments in each age group from 2002 to 2013, where 2002 serves as the benchmark year:[1]

$$\text{R-share}_{i,t} = \sum_{t=2003}^{2013} \beta_t \times Year_t + FE_{NAICS3} + \epsilon_{i,t}, \tag{2}$$

Panels A to D in Figure 1 depict the year-specific coefficients β_t for each establishment type (also shown in Table IA.1). They look strikingly similar: the decline in R-share is comparable for all three establishment types. In the subsequent sections, we explore in more detail the characteristics specific to each type.

Figure 1: **Evolution of Routine Share**



Notes: Panels A-D plot establishments' R-share evolution for all establishment types (see equation (2)). Vertical bars represent the point estimates' robust standard errors.

[1]Throughout the paper, all observations are weighted by the product of the establishment's total employment and the BLS sampling weight and we report robust standard errors.

2.1 Incumbents

As a first step, we regress an incumbent establishment's R-share on its age and establishment fixed effects,

$$\text{R-share}_{i,t} = \sum_{j} \gamma_j AgeGroup_j + FE_{Est} + \epsilon_{i,t}, \tag{3}$$

where $AgeGroup$ specifies seven 3-year age bins. γ_j captures an establishment's R-share as it ages, with the [0-2] age group serving as the baseline. Adding establishment fixed effects allows us to disentangle how much of the R-share evolution comes from an establishment's life cycle relative to time.

Table 1's first column reveals that R-share declines with age: compared to its initial level, an establishment's R-share drops by 2.4 ppt by ages [12-14] and 4 ppts by age 20. All differences are significant at the 1% level. Moreover, this conclusion holds even if we focus on different establishment cohorts, as shown in Columns (2)-(5).

Finally, considering that firms grow with age, we control for an establishment's or parent firm's size in columns (6) and (7). The age coefficients remain unaffected, confirming the life cycle dimension of R-share dynamics.

2.2 Entrants and exiters

Next, we investigate whether creative destruction, through entry and/or exit, plays a significant role in the decline of the aggregate R-share. We start by running the following regression:

$$\text{R-share}_{i,t} = \theta_1 E_{i,t} + FE_{NAICS3 \times Year} + \epsilon_{i,t}, \tag{4}$$

where $E \in \{Entrant, Exiter\}$ is a dummy variable that equals one if establishment i is an entrant/exiter in year t. Industry interacted with time fixed effects are included.

Column (1) of Table 2 reports the coefficient θ_1 when we only include the entrant dummy. The result shows that, on average, firms at entry are characterized by a *higher* R-share than their incumbent industry peers, a difference of 0.34 ppt. This casts doubt on the contribution of entrants through a creative destruction channel, whereas they would be more likely to enter with newer technologies (and have lower R-share) than incumbents. On the other hand, Column (2) shows that in the exiter's last year of existence, the R-share was 0.44 ppt higher than that of its peers. Column (3) confirms similar outcomes when both dummies are included.[2]

[2]Focusing only on establishments with over 20 employees does not change this result (see Table IA.2).

Table 1: **Within Incumbents R-Share**

	All	1990 Cohort	1995 Cohort	2000 Cohort	2005 Cohort	All	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age[3-5]	-0.623***			-0.692***	-0.539***	-0.623***	-0.625***
	(0.105)			(0.166)	(0.147)	(0.105)	(0.105)
Age[6-8]	-1.235***		-0.544**	-1.343***	-1.135***	-1.234***	-1.237***
	(0.123)		(0.229)	(0.173)	(0.225)	(0.124)	(0.123)
Age[9-11]	-1.914***	-0.826	-1.333***	-1.861***		-1.912***	-1.917***
	(0.142)	(0.809)	(0.229)	(0.195)		(0.143)	(0.142)
Age[12-14]	-2.414***	-1.199	-1.806***	-2.797***		-2.413***	-2.419***
	(0.163)	(0.806)	(0.237)	(0.362)		(0.163)	(0.163)
Age[15-17]	-2.975***	-1.926**	-2.183***			-2.974***	-2.978***
	(0.184)	(0.828)	(0.266)			(0.184)	(0.184)
Age[18-20]	-3.658***	-2.604***	-2.029***			-3.657***	-3.660***
	(0.211)	(0.815)	(0.582)			(0.211)	(0.211)
Age[21-22]	-3.998***	-2.897***				-3.997***	-4.001***
	(0.310)	(0.850)				(0.310)	(0.310)
Log(Emp)						-0.015	0.064
						(0.110)	(0.058)
N	1,280,804	292,775	363,542	402,364	212,605	1,280,804	1,280,804
R^2	0.91	0.91	0.91	0.90	0.92	0.91	0.91

Notes: Results of regressing establishment routine share on its age with establishment fixed effects. The benchmark age for all columns is Age[0-2], except for Columns (2) and (3) where the benchmark is Age[6-8] and Age[3-5], respectively. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table 2: **R-share: Entrants & Exiters**

	(1)	(2)	(3)
Entrant	0.335***		0.310***
	(0.083)		(0.083)
Exiter		0.439***	0.412***
		(0.127)	(0.128)
N	3,032,548	3,010,740	3,010,740
R^2	0.66	0.66	0.66

Notes: Results of regressing establishment routine share on entrant and exiter dummy with NAICS3-Year fixed effects. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Next, we study R-share dynamics *around* entry or exit. First, for each entering establishment in period t, we track its R-share in $t+1, ..., t+7$ relative to its incumbent peers by estimating the following regression:

$$\text{R-share}_{i,t} = \lambda_\tau^E Entrant(-\tau)_{i,t} + FE_{NAICS3-Year} + \epsilon_{i,t}, \tag{5}$$

where $Entrant(-\tau)_{i,t}$ is equal to one if establishment i was an entrant τ years ago, and zero otherwise, up to $\tau = 7$. We exclude establishments younger than τ.

Panel A of Figure 2 depicts λ_τ^E. We find no evidence that following entry, new establishments on average ever display a lower R-share than their peers. This confirms that the entry margin does not contribute to the overall decline in R-share.[3]

Next, we turn to the dynamics prior to exit. For each exiting establishment in period t, we estimate its R-share in $t-7, ..., t-1$ relative to incumbents by running the following regression:

$$\text{R-share}_{i,t} = \lambda_\tau^X Exiter(\tau)_{i,t} + FE_{NAICS3-Year} + \epsilon_{i,t}, \tag{6}$$

where $Exiter(\tau)_{i,t}$ is a dummy variable that is equal to one if establishment i will be an exiter in τ years, and zero if the establishment will survive beyond τ years.

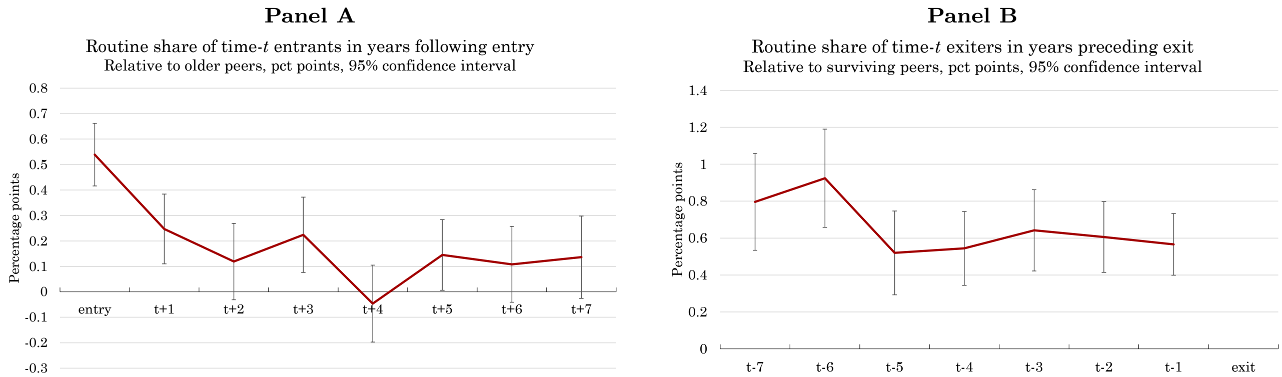
Panel B of Figure 2 depicts the coefficients λ_τ^X. We find that exiters had a significantly higher R-share than their incumbent peers many years before their eventual death. Hence, exiters fall behind their surviving peers' R-share evolution years before exit, and are not simply the victims of an exit-inducing shock.

2.3 A decomposition of the evolution of the routine share

Finally, we show in Table 3 the result of a Melitz-Polanec decomposition of the within-industry R-share. We find that 1.74 ppt out of the total 1.99 ppt fall in the R-share is coming from the within-

[3]Focusing only on entrants that never exit later in the sample does not change this result (see Figure IA.1).

Figure 2: **Routine Share Dynamics of Entrants and Exiters**



Notes: Panel A plots entrants' R-share relative to their existing incumbent peers. Panel B plots exiters' R-share relative to their surviving incumbent peers.

incumbent margin, with another 1.03 from shifts in weights across incumbents. The contribution of entry, at 0.69, is *positive*, in line with our earlier results. That of exit, while negative, is very small (-0.01). The cross-term rounds out the total, at 0.1.

All in all, the decomposition confirms that incumbent establishments are the main drivers of the decline in the U.S. R-share.

Table 3: **Decomposition of Routine Share Change**

Total	Within	Chg.Weight	Cross-Term	Net Entry	Net Entry	
					Entry	Exit
-1.99	-1.74	-1.03	0.10	0.68	0.69	0.01

3 Conclusions

Over time, entrants, exiters and incumbents all exhibit a reduction in their routine employment share. The driving factor for the R-share fall is a decline in incumbents' R-share as they age. Thus, research on R-share reduction should focus on occupational dynamics within existing establishment.

References

Acemoglu, Daron and David H. Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 2011, *4.*

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, "The Polarization of the U.S. Labor Market," *American Economic Review*, May 2006, *96.*

Bockerman, Petri, Seppo Laaksonen, and Jari Vainiomaki, "Does ICT Usage Erode Routine Occupations at the Firm Level?," *Labour*, 2019, *33.*

Goos, Maarten and Alan Manning, "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *Review of Economics and Statistics*, February 2007, *89.*

Harrigan, James, Ariell Reshef, and Farid Toubal, "The March of the Techies: Job Polarization Within and Between Firms," *Research Policy*, 2021, *50.*

Heyman, Fredrik, "Job polarization, job tasks and the role of firms," *Economics Letters*, 2016, *145.*

Jaimovich, Nir and Henry E. Siu, "Job Polarization and Jobless Recoveries," *Review of Economics and Statistics*, March 2020, *102.*

Online Appendix for

Under the Hood of Routine Share Decline

Nir Jaimovich, Miao Ben Zhang, and Nicolas Vincent

A Details on Data and Measures

A.1 BLS Microdata

We use two sets of confidential microdata from the Bureau of Labor Statistics (BLS) in this study. The first one is the BLS Occupational Employment and Wage Statistics (OEWS) microdata, which is a stratified sample of about 1.2 million establishments from a universe of approximately 6.8 million non-farm establishments from the Quarterly Census of Employment and Wages (QCEW) database.[4] The QCEW database is provided to the BLS by state workforce agencies that collect unemployment insurance (UI) reports from employers. Employers are required by law to file these reports to the state where each establishment is located. The establishments in the OEWS sample frame are stratified by establishments' industry, geography, and size, and each establishment is given a sampling weight by the OEWS.

Out of the 1.2 million establishments in the sample frame, the OEWS program surveys about 400,000 establishments each year, with one establishment surveyed in every 3 years to reduce the response burden. That is, an establishment in the sample framework is surveyed in years $t-3$ and t but not in between. From each OEWS survey, we obtain the establishment's number of workers in each occupation and wage bin, whereas occupation is defined based on over 800 categories, and wage bin is specified by the BLS based on about 12 bins of hourly wages.[5] From the survey results, we obtain an establishment's total number of employees in each occupation.

The OEWS survey maintained a consistent industry classification (based on the NAICS) and occupation classification (based on the SOC) after 2002. The BLS surveys about 200,000 establishments in May and November of each year and reports aggregate statistics in May of each year by weighting establishments in the previous six surveys. Because the establishment weights are assigned in May of each year, we thus regard survey results from November of year t and May of year $t+1$ as the

[4]BLS uses this microdata to produce aggregate occupation statistics at https://www.bls.gov/oes/tables.htm. The employees in the covered establishments represent 62% of the U.S. The survey covers all industries except for agricultural workers, private households, and unincorporated self-employed workers without employees.

[5]See a recent form of the OEWS survey at https://www.bls.gov/respondents/oes/pdf/forms/uuuuuu_fillable.pdf. See more details of the survey methods and statements for each year from the BLS documentation archives at https://www.bls.gov/oes/oes_doc_arch.htm

observations for the year ends at t. For instance, our last year 2013's sample includes surveys done in November 2013 and May 2014, where May 2014 is the last period of the microdata we accessed. Our OEWS sample thus includes establishments from 2002 to 2013. For each establishment, we have information about the number of employees and average hourly wage rate per employee in each occupation, as well as the establishment's unique identifier, sampling weight, employer identification number (EIN), government ownership, county code, and industry code.

The second microdata we accessed is the establishment identifiers of the QCEW universe for all but ten states from 1990 to 2014.[6] This microdata is helpful for us to identify the entry and exit years of each establishment in our OEWS sample. Specifically, we define the first year of the establishment in the QCEW universe as the entry year of the establishment and the last year in the QCEW universe as the exit year. For establishments that exist in the QCEW universe before 1990, we do not know their precious entry year, and we treat them as incumbents throughout our sample period.

Merging the QCEW and the OEWS microdata results in our final sample of 24.2 million establishment-occupation-year observations for non-government-owned establishments with the QCEW information (i.e., from states outside the ten states), covering about 257,000 establishments each year from 2002 to 2013. In all of our analyses, we weigh each establishment using the product of the establishment's total employment and the BLS sampling weight.

A.2 Measuring Routine Occupations

We follow the definition of Jaimovich and Siu (2020) and categorize occupations into three categories. In particular, occupations are regarded as **routine (R)** if they are "sales and related occupations (SOC2=41)," "office and administrative support occupations (SOC2=42)," "production occupations (SOC2=51)," "transportation and material moving occupations (SOC2=53)," "construction and extraction occupations (SOC2=47)," and "installation, maintenance, and repair occupations (SOC2=49)." Non-routine occupations are further divided into **non-routine cognitive (C)** if they are "management, business, and financial operations occupations (SOC2=11)" and "professional and related occupations (from SOC2=13 to SOC2=29)," and **non-routine manual (M)** if they are "service occupations (from SOC2=31 to SOC2=39)."

B Additional Results

Table IA.1 presents the detailed regression results behind the panels of Figure 1. They show a distinct fall in the level of the R-share over time for incumbents, entrants, and exiters alike.

[6] The ten states that we do not have the QCEW data include Florida, Kentucky, Massachusetts, Michigan, Mississippi, Missouri, New Hampshire, New York, Pennsylvania, and Virginia.

Table IA.1: **Evolution of Routine Share Over Time**

	All	Incumbent	Entrant	Exiter
Year=2003	-0.400**	-0.395**	-0.163	-1.044
	(0.170)	(0.186)	(0.405)	(0.652)
Year=2004	0.181	0.458***	-1.136***	-0.594
	(0.159)	(0.174)	(0.384)	(0.609)
Year=2005	-0.149	-0.025	-0.587	-1.678***
	(0.162)	(0.178)	(0.381)	(0.595)
Year=2006	-0.759***	-0.626***	-1.193***	-1.794***
	(0.175)	(0.194)	(0.374)	(0.604)
Year=2007	-0.958***	-0.812***	-1.494***	-1.793***
	(0.169)	(0.186)	(0.378)	(0.607)
Year=2008	-1.328***	-1.197***	-1.724***	-1.548***
	(0.170)	(0.186)	(0.398)	(0.596)
Year=2009	-1.999***	-1.828***	-2.689***	-2.532***
	(0.181)	(0.200)	(0.394)	(0.645)
Year=2010	-1.764***	-1.608***	-2.314***	-2.434***
	(0.169)	(0.185)	(0.403)	(0.611)
Year=2011	-1.835***	-1.653***	-2.665***	-1.950***
	(0.166)	(0.181)	(0.424)	(0.616)
Year=2012	-2.112***	-2.005***	-2.252***	-1.726**
	(0.182)	(0.199)	(0.413)	(0.745)
Year=2013	-2.368***	-2.275***	-2.523***	-2.111**
	(0.171)	(0.186)	(0.424)	(0.957)
N	3,032,551	2,532,651	499,899	245,918
R^2	0.66	0.67	0.62	0.59

Notes: This table reports the results of regressing an establishment's share of routine-task labor, in percentage, on year dummies while controlling for industry fixed effects at the NAICS3 level (see equation (2)). All regressions are weighted by a product of establishment employment and the sampling weight of the establishment assigned by the BLS. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period is from 2002 to 2013, where 2002 serves as the benchmark year.

Table IA.2 provides additional results for the entry/exit regressions. In particular, we restrict the sample to larger establishments in Column (4) and control for establishment size in Column (5). See table notes for details.

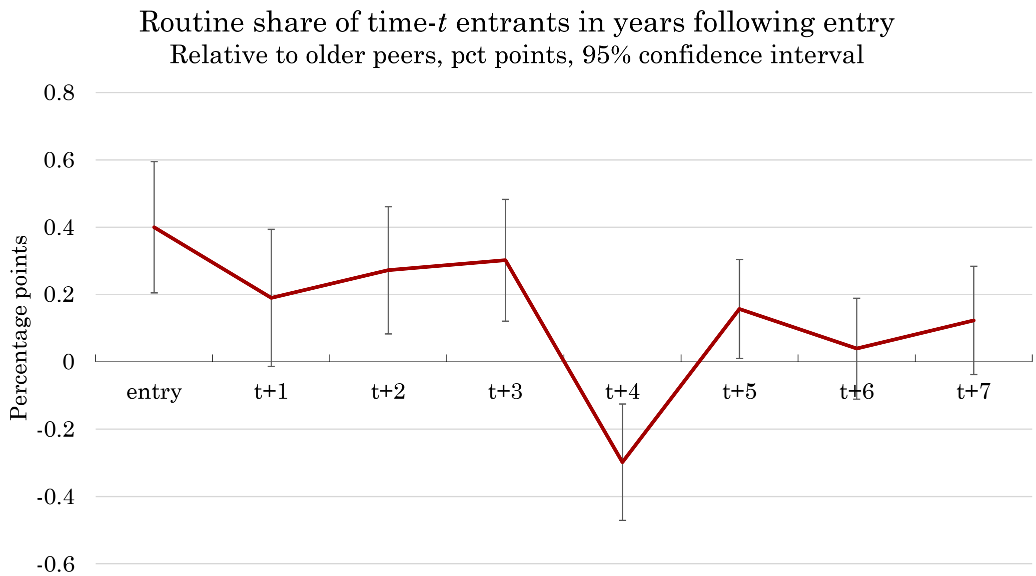
Table IA.2: **Additional Robustness Results for R-Share in Entrants & Exiters**

	(1)	(2)	(3)	(4)	(5)
Entrant	0.335***		0.310***	0.604***	0.511***
	(0.083)		(0.083)	(0.112)	(0.106)
Exiter		0.439***	0.412***	0.764***	0.697***
		(0.127)	(0.128)	(0.163)	(0.153)
N	3,032,548	3,010,740	3,010,740	1,264,412	1,264,339
R^2	0.66	0.66	0.66	0.69	0.71

Notes: This table reports the results of regressing establishment routine share on an entrant dummy and an exiter dummy with NAICS3-Year fixed effects (see equation (4)). Column (4) restricts the sample to establishments with at least 20 employees, while Column (5) further includes NAICS3-Year-EmpBin fixed effects where the establishment employment bins are [20-49], [50-99], [100-249], [250-499], and [500+]. All regressions are weighted by the product of the establishment's total employment and the BLS sampling weight. Robust standard errors are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Figure IA.1 depicts the dynamics of entrants following birth, focusing only on entering establishments that never exit afterward. This allows for ruling out a role for post-entry exit through selection.

Figure IA.1: **Robustness: The Routine Share Dynamics of Entrants Among Survivors**



Notes: This figure plots Panel A of Figure 2 using establishments that survive at $t+7$. The vertical bars represent the robust standard errors of the point estimates.