Discussion Paper Series

IZA DP No. 18826 July 2026

The Anatomy of Polarization: Evidence from Worker Flows

Fabio Cerina
University of Cagliari and CRENoS

Elisa Dienesch
University of Cagliari, CRENoS and AMSE

Alexander Monge-Naranjo
Federal Reserve Bank of Atlanta, Emory University and CEPR

Alessio Moro
University of Cagliari, CEPR and IZA@LISER

IZA@LISER NETWORK
ISSN: 2365-9793

The Anatomy of Polarization: Evidence from Worker Flows*

Abstract

Using a French administrative panel following workers over 1984–2021, we quantify how worker flows account for employment polarization. We trace job-to-job flows across manual, routine, and abstract occupations, and flows between employment and non-employment. Polarization emerges after 1994 almost entirely through a shift in the latter: a collapse in entry into routine jobs, concentrated in routine-cognitive occupations. Job-to-job upgrading from routine to abstract remains large and stable, with non-college workers accounting for 76% of net flows. Manual employment grows mainly by absorbing entrants. Polarization reflects sorting at entry and re-entry rather than incumbent displacement, motivating life-cycle models incorporating both margins.

JEL classification

J62, J24, J21

Keywords

labor market polarization, worker flows, occupational mobility, routine-biased technological change, longitudinal data

Corresponding author

Alessio Moro

amoro@unica.it

* We thank participants of the workshops in Cagliari (Brucchi Luchino; "Inequalities and Labor Markets"), of the Conference in honor of Chris Pissarides in Cyprus and of the seminars in Aix-en-Provence. We thank Pedro Gomes, Robert Shimer and Etienne Wasmer, for useful suggestions. We also thank Simone Nobili for excellent research assistance. Cerina, Dienesch and Moro acknowledge financial support from the Italian PNRR funded by NextGenerationEU «The Role of Skill-Biased Technological Change in Immiserizing Growth, Spatial Polarization, and Product Quality Upgrading» (CUP F53D23003240006); Monge-Naranjo acknowledges financial support from the Research Council of the European University Institute. The views expressed are the authors' alone and do not reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. All errors are our own.

1 Introduction

Employment polarization is one of the most consequential labor-market transformations in advanced economies of the past four decades: middle-skill, routine-intensive occupations have contracted, while high-skill abstract and, to a lesser extent, low-skill manual jobs have expanded (Autor et al., 2003, 2006; Goos et al., 2014; Cerina et al., 2021, 2023). Virtually all of this evidence comes from repeated cross sections. Cross sections record changes in occupational shares but are silent about the worker flows behind them, and radically different reallocation processes can be consistent with the same aggregate pattern. For instance, the decline of routine employment may reflect the downgrading of incumbent routine workers into manual jobs and non-employment (as in Autor and Dorn, 2013). Alternatively, it may reflect entrants who sort away from routine work while incumbents keep moving up the occupational ladder. These two reallocation processes involve different workers, imply different distributional and welfare consequences, and call for different policy responses. Distinguishing between them requires observing worker flows rather than stocks at different points in time, and flows require longitudinal data. Previous work in this literature studied these flows for the United States (Cortes, 2016; Cortes et al., 2017, 2020), but the available panels combine small samples followed at low frequency with survey-reported occupations, or match workers only over short windows.

This paper exploits exhaustive French administrative data that track individual careers for nearly four decades to quantify how much each margin (job-to-job flows, entry, and exit) accounts for the observed employment polarization. We characterize workers' reallocation using a longitudinal matched employer–employee panel that follows individual workers in France from 1984 to 2021. The panel covers all salaried employment (*tous salariés*) and records each worker's annual occupation and non-employment spells. We classify occupations into three broad groups—low-skill manual (m), middle-skill routine (r), and high-skill abstract (a)—and we track transitions among these categories and in and out of non-employment (o). From these records we construct annual transition matrices and decompose year-to-year changes in occupational shares into (i) net job-to-job flows among the three occupational categories and (ii) net flows between each occupation and non-employment. This decomposition answers three questions about the emergence of polarization in France that cross sections cannot: whether the workers most affected are entrants or incumbents; whether career advancement across occupational categories has accelerated or stalled; and whether polarization emerges through downgrading, through a failure of entry to replenish middle-skill jobs, or through enhanced upward mobility from the bottom to the top of the occupational ladder.

The behavior of the French labor market changes sharply in the mid-1990s. Before 1994, aggregate employment shares show no clear trend and the routine share is broadly flat. After 1994, the routine share declines steadily while abstract and manual employment expand. We therefore split the sample into a pre-polarization phase (1984–1993) and a polarization phase (1994–2021), a break that coincides with the sharp fall in information and communication technology (ICT) costs documented in the literature, and we identify which flows changed across the two regimes. Our central finding is that the cross-sectional appearance of polarization is not driven by widespread downward displacement of routine incumbents. It is driven by the margins connecting employment to non-employment, and in particular a collapse in the flow of workers entering routine occupations from non-employment. We unpack this result in two steps below.

Figure 1 shows the key changes in the net flows of workers across occupational categories and in and out of non-employment. The first key finding is the prominent change in and out of non-employment flows, most notably the sign flip for routine occupations beginning around 1994. Before 1994 non-employment is a net feeder into routine: net flows from non-employment to routine add roughly +0.25 percentage points per year to the routine share, helping to keep the aggregate routine share flat. After 1994 this channel reverses: net non-employment flows drain routine, subtracting roughly 0.1 percentage points per year over 1994–2021, a total change of about −0.35 percentage points. This reversal reflects a collapse in both entry rates from non-employment into routine and exit rates from routine into non-employment, of comparable proportional magnitude. The entry-side fall dominates in absolute terms flipping the net flow. The exit-side contraction partially offsets the routine share decline *rather than driving it.* New entrants flow into routine occupations at a substantially slower rate. Combined with a persistent routine→abstract upgrading drain, this entry-margin collapse triggers the aggregate routine decline.

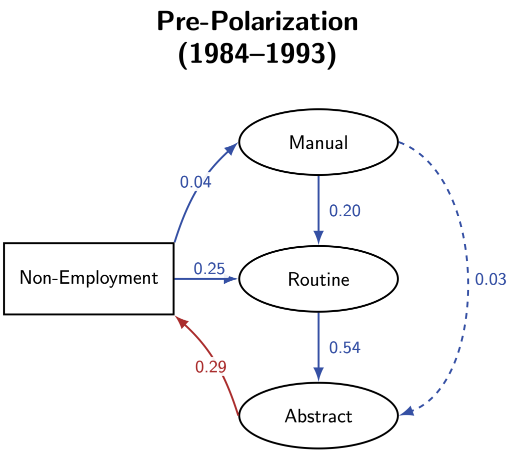

Figure 1: Net yearly worker flows in the pre-polarization (1984–1993) and polarization (1994–2021) regimes in percentage points. Blue arrows mark upgrading and entry-from-non-employment. Red arrows mark downgrading or outflows to non-employment.

Within routine itself, this entry-margin collapse is concentrated in the cognitive component, that is, clerical, administrative, and supervisory occupations. Net non-employment flows into *routine-cognitive* work add roughly +0.80 percentage points per year to its share before 1994 and drop to +0.04 in the polarization era, a swing of about −0.76 percentage points. The corresponding flows into *routine-manual* work move in the opposite direction, with their drain actually weakening over the same span. The 1994 sign change in the non-employment channel is mainly a routine-cognitive phenomenon, the routine type most directly displaceable by ICT equipment.

A second key result from the changes in the flows is the absence of net downgrading patterns, even in the polarization period. Job-to-job flows from routine to abstract are large, persistent, and stable across the panel at 0.52–0.54 percentage points per year. The net flow from manual into routine remains positive throughout the period, confirming the upgrading trend, though its decline after 1994 contributes to the fall in routine share. The rise in the manual share is driven entirely by the non-employment channel: manual occupations absorb a disproportionate share of entrants from non-employment (+0.25 pp/year), while net flows to the other two types of occupations actually subtract from the manual share (−0.13 pp/year). These facts are inconsistent with a displacement interpretation of routine-biased technical change. In the French microdata the displacement channel is absent.

A third finding is that the changes in employment shares across occupational categories are remarkably different across education groups. Among non-college workers the typical U-shape emerges: routine falls by about 13 percentage points over 1984–2021 while both manual and abstract rise (+7.7 and +5.7 pp respectively).[1] Almost half (43%) of the 11.5-percentage-point rise in the abstract share over 1984–2021 is driven by non-college workers. The rising college share contributes roughly 50%, with within-college upgrading making up the remaining 7%. The non-college contribution to routine-abstract upgrading *flows* is even higher. Specifically, non-college workers account for 76% of all net routine-to-abstract worker-flows in both periods. The displacement view according to which non-college routine workers are forced into manual employment (Autor and Dorn, 2013), is thus not confirmed by French worker flows. The evidence suggests a more optimistic version of the routinization hypothesis, at least for workers already in routine occupations: a persistent ICT-complementary abstract demand sustained a stable upgrading channel for non-college routine workers, so that the collapse of routine demand reallocated entrants away from routine rather than displacing incumbents downward into manual employment. College-educated workers, by contrast, exhibit essentially flat employment shares in the polarization era: their routine, abstract, and manual shares each change by no more than 1.5 percentage points between 1994 and 2021.

Worker-level conditional analysis reinforces these patterns. Education is the dominant predictor of upgrading and downgrading: holding age, gender, and tenure fixed, college

[1] In a companion paper on the spatial dimension of polarization in France, Cerina et al. (2026b) document a complementary pattern: the routine-to-abstract upgrading channel runs at roughly twice the intensity in large cities as in small ones.

roughly triples the routine→abstract upgrading rate and substantially reduces downgrading from abstract. Occupation-specific tenure stabilizes workers, so polarization runs mainly through entrants and short-tenured workers while women face an additional disadvantage on upgrading margins.

Our findings connect to the task-based/automation literature on the sources of polarization (Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2020; Acemoglu and Loebbing, 2026) and, most directly, to the line of research that first placed worker flows at the center of the analysis of routine employment. Using the PSID, Cortes (2016) follows U.S. routine workers over time and documents strong selection in who leaves routine work: switchers into abstract jobs come from the top of the routine wage distribution, switchers into manual jobs from the bottom. Cortes et al. (2017) show that the per-capita decline of U.S. routine employment is accounted for primarily by changes in the rates at which workers flow into and out of routine work—above all, falling inflows from non-employment concentrated among young and less-educated workers—rather than by shifts in demographic composition. Jaimovich and Siu (2020) show that these routine losses bunch in recessions, linking polarization to jobless recoveries. Closest to our approach, Cortes et al. (2020) use matched CPS data to decompose the U.S. routine employment decline and conclude that it is driven primarily by a collapse in non-employment-to-routine inflow rates rather than by rising routine-to-non-employment exit rates. That we reach the same verdict on the entry margin for France indicates that the mechanism is not specific to the U.S. labor market. These authors pursued the flows question with the data available for the United States—household surveys whose small samples or short rotation windows limit how finely individual transitions can be tracked. Our French DADS–EDP panel, which follows individual workers continuously across four decades, complements their evidence by recovering the job-to-job margin as well, and by showing that it is dominated by routine-to-abstract upgrading rather than routine-to-manual displacement. We are not aware of previous work characterizing employment polarization through worker flows at this level of detail.

Relatedly, Charlot et al. (2024) use the quarterly French Labor Force Survey over 2003–2019 to study how non-standard contracts shape routine employment dynamics over the business cycle, and find that inflows into routine standard work – from routine non-standard work and from unemployment – are the main drivers of its cyclical variation. We interpret their cyclical inflow-margin finding as the business-cycle counterpart of our long-run approach.[2]

The upgrading pattern we document is also consistent with the stepping-stone view of routine employment proposed by García-Peñalosa et al. (2023), in which routine cognitive jobs serve as a human capital ladder into abstract work rather than as an absorbing

[2]Our flows are measured annually rather than quarterly, so a single transition may bundle some within-year moves. Annual and quarterly measures therefore need not coincide.

terminal state.[3] Bocquet (2024) also uses French DADS data to study occupational reallocation, modeling it as a network of feasible transitions with bridge occupations as bottlenecks; his main research question is why reallocation is slow, whereas ours is which margins drive polarization in the first place.

Methodologically, our contribution is to integrate entry into and exit from non-employment into the flow decomposition of polarization and to provide a detailed accounting of who makes which transitions. Substantively, the facts motivate quantitative life-cycle models of occupational choice that allow for cohort and entry-margin effects and that can evaluate the distributional and welfare consequences of polarization on top of different triggers of it.

The remainder of the paper proceeds as follows. Section 2 describes the DADS-EDP administrative panel and our sample construction. Section 3 develops the accounting framework. Section 4 documents the timing and magnitudes of flow changes and reports robustness checks. Section 5 examines heterogeneous drivers of polarization. Section 6 concludes.

2 Data

2.1 The DADS-EDP Panel

We use the Déclaration Annuelle des Données Sociales — Échantillon Démographique Permanent (DADS-EDP), a longitudinal matched employer–employee administrative panel. After sample restrictions the panel covers roughly 1.9 million workers over 1984–2021 (1990 is absent from the DADS series). DADS-EDP links annual DADS wage records to the Permanent Demographic Sample (EDP), a birth-date stratified sample of the French population.[4] Coverage includes both private and public salaried employment (*tous salariés*).

Each worker is observed annually. For every employment spell within a year the data record the two-digit PCS occupation code (CS2), employer location (*aire urbaine*), and gross and net wages.[5] We assign a main occupation as the spell with the largest share

[3]A recent working paper along these lines is Kashkarov and Artemev (2024), who combine PSID data with historical vacancy records to argue that a large share of U.S. high-skill workers transit through routine cognitive jobs on their way to abstract occupations. Our French evidence is qualitatively consistent with this narrative, with the added qualifier that the stepping stone is active primarily for college-educated workers.

[4]Prior to 2002 the sample covered workers born in October of even-numbered years (about 1/24 of the population); the 2002 expansion added workers born in October of odd-numbered years, roughly doubling the sample.

[5]For 156,389 worker-years (126,520 unique workers, 6.7% of the panel) the raw CS2 is missing in DADS–EDP while wage and employment are observed; these are predominantly civil servants of the *fonction publique d'État* whose administrative classification did not flow into DADS before the Grand Format integration finalised in 2009. We recover the missing codes by carrying the worker's own observed CS2 from the temporally nearest panel year (97.8% of cases, with a median temporal distance of one

of hours in that year. Other variables include birth year, gender, and panel entry year. Education is recorded in the EDP from 2002 onward when the panel was linked to census records; for workers observed both before and after 2002 we extend the education field backward to pre-2002 spells. After this extension education is available for 1,427,753 workers (75.3% of the panel). We define a binary college indicator equal to one for workers holding a *Licence* or higher degree.

The panel experienced two major expansions. In 2002 the addition of about 845,000 mid-career workers (median age 36) more than doubled the sample; in 2006 a smaller expansion added 181,000 workers (median age 24). Both expansions create year-to-year discontinuities that we address explicitly in the flow analyses (Section 2.3).

2.2 Occupation Classification

We group occupations into three broad categories based on CS2 codes, plus a non-employed state:

State	**Label**	**CS2 codes**
m (Manual)	Low-skill service & manual	10, 53, 55, 56, 64, 68, 69
r (Routine)	Clerical, supervisory, production	42, 43, 45, 46, 48, 52, 54, 62, 63, 65, 67
a (Abstract)	Managerial, professional, technical	23, 31, 33, 34, 35, 37, 38, 47
o (NE)	Non-employed	Not in salaried employment

The classification is task-based rather than wage-ranked[6] but the mean wage ranks are stable over time: Abstract earns roughly 3× the Manual average and Routine about 1.5×, and the ordering $a > r > m$ holds in every benchmark year (Appendix A.1, Table 6).

2.3 Sample Restrictions and Corrections

We restrict the sample to workers aged 25–64, the standard prime working-age window; this excludes most apprentices and trainees and those past statutory retirement age.

Two data features require special treatment in the flow decomposition. First, the 2002 and 2006 panel expansions introduce discontinuities in year-to-year flow counts; accordingly, the transition pairs $(2001, 2002)$ and $(2005, 2006)$ are excluded from the baseline analysis. Second, 1990 is missing from DADS, so the pair $(1989, 1991)$ spans two years rather than one; we include this pair throughout but halve its flows to obtain annual equivalents.

year). Details and a no-imputation robustness check are in Appendix C.4.

[6] For instance, our classification slightly differs from the wage-rank scheme of Davis et al. (2026) in two CS2 codes: CS2=47 (technicians) is Abstract here but middle-paid there, and CS2=64 (drivers) is Manual here but middle-paid there. All qualitative results are robust to reclassifying these codes.

3 Accounting Framework

This section develops the accounting identities that link worker-level flows to the evolution of occupational employment shares and non-employment.

3.1 Setup and Notation

Following the classification in the previous section, we group workers into one of four states: three employed occupations — manual (m), routine (r), and abstract (a) — and an 'out-of-employment' state (o). The o state covers all individuals in the panel not observed in private- or public-sector salaried employment in a given year. We use the subindex $k \in \{m, r, a\}$ to denote an employed occupation and write o explicitly for non-employment; where an index ranging over all four states is needed we use $\kappa \in \{m, r, a, o\}$. We use e for the 'employment' state, i.e., the aggregate of workers in $k \in \{m, r, a\}$.[7]

We measure transition flows between consecutive years $t-1$ and t for workers aged 25–64 at $t-1$. Employment levels $N_k(t)$ are taken from cross-sectional snapshots of the 25–64 window at each date.[8]

Let $N_e(t) \equiv N_m(t) + N_r(t) + N_a(t)$ denote total employment in year t. The *employment share* of occupation k is

$$E_k(t) \equiv \frac{N_k(t)}{N_e(t)}, \quad k \in \{m, r, a\}. \tag{1}$$

These shares sum to one: $E_m + E_r + E_a = 1$.

We also consider employment shares that include the out-of-employment workers in the panel. We define *augmented shares* using the French National Statistical Institute (INSEE) employment rate $e(t)$ for the 25–64 age group:[9]

$$S_k(t) = E_k(t) \cdot e(t), \quad k \in \{m, r, a\}, \qquad \text{and} \qquad S_o(t) = 1 - e(t). \tag{2}$$

These four shares sum to one: $S_m(t) + S_r(t) + S_a(t) + S_o(t) = 1$. The total working-age

[7]DADS-EDP captures only salaried (and assimilated-salaried) employment. Some categories of independent employment outside its coverage are those exercising under non-salaried status: artisans, commerçants and chefs d'entreprise registered as *travailleurs indépendants* with URSSAF; some liberal professions practicing in own cabinet (e.g., *médecins libéraux, infirmiers libéraux)*; and agriculteurs exploitants. Salaried workers in the same professions — e.g., hospital-employed doctors and nurses, lawyers employed by firms, accountants in salaried positions — *are* in DADS-EDP. Strictly speaking, the o state is therefore more accurately interpreted as *non-observed-employment* rather than *non-employment*: it pools genuine non-employment with these uncovered *non-salaried* categories. The implications of this pooling for the analysis are explored in detail in Appendix C.3. Under the data-anchored bias correction implemented in that appendix, the qualitative findings of the paper are preserved and, if anything, *reinforced*.

[8]So workers must be aged 25–64 at $t-1$ for the $t-1$ stock and aged 25–64 at t for the t stock. Boundary cohorts (aging-in at 24 $\rightarrow$ 25, aging-out at 64 $\rightarrow$ 65) are handled by the residual recovery of $F_{o,k}(t)$ in Section 3.2; the coding convention is in Appendix A.3.

[9]The employment-rate series is constructed as a uniform-age-year-weighted average of the employment rates $e_{25\text{–}49}(t)$ and $e_{50\text{–}64}(t)$, the two sub-group rows in INSEE table T207: $e(t) = \frac{25}{40}\, e_{25\text{–}49}(t) + \frac{15}{40}\, e_{50\text{–}64}(t)$.

population is then $N(t) = N_e(t)/e(t)$, and the stock of non-employed individuals aged 25–64 is recovered as

$$N_o(t) \;=\; N(t) - N_e(t) \;=\; N_e(t) \cdot \frac{1 - e(t)}{e(t)}. \tag{3}$$

Using the INSEE rate avoids relying on the DADS stock of non-employed individuals, which is less well identified in the panel than the employment stocks by occupation.

3.2 Stocks, Flows, and the Employment-Share Changes

We now consider the worker flows across the four states. Let $F_{\kappa,\kappa'}(t)$ denote the number of workers in state κ at $t-1$ and in state κ' at t, with $\kappa, \kappa' \in \{m, r, a, o\}$. The 4×4 flow matrix $[F_{\kappa,\kappa'}]$ has three components of substantive interest: **(i)** *job-to-job transitions* (JtJ), $F_{k,k'}$ for $k, k' \in \{m, r, a\}$ with $k \neq k'$; **(ii)** *exits from employment*, $F_{k,o}$ for $k \in \{m, r, a\}$; and **(iii)** *entries into employment*, $F_{o,k}$ for $k \in \{m, r, a\}$.

Components (i) and (ii) are directly observed in our panel. Components (iii) are harder to identify, because the DADS panel does not reliably distinguish genuine non-employment from absence from the panel before a worker's first observation.

The net job-to-job inflow to occupation k from the other two employed occupations is

$$\text{Net\_JtJ}_k(t) \;\equiv \sum_{k' \in \{m,r,a\},\, k' \neq k} \Big[F_{k',k}(t) - F_{k,k'}(t)\Big]. \tag{4}$$

Employment stocks and worker flows are then linked by the accounting identity: for each occupation $k \in \{m, r, a\}$ and all t, the change in k's stock equals its net job-to-job inflow plus its net flow with non-employment,

$$N_k(t) - N_k(t-1) \;=\; \text{Net\_JtJ}_k(t) \;+\; \Big[F_{o,k}(t) - F_{k,o}(t)\Big]. \tag{5}$$

Job-to-job flows and exits $F_{k,o}$ are observed in the matched panel, but entries $F_{o,k}$ are not; we therefore solve the identity for the entries,

$$F_{o,k}(t) \;=\; \Delta N_k(t) + F_{k,o}(t) - \text{Net\_JtJ}_k(t), \tag{6}$$

where $\Delta N_k(t) \equiv N_k(t) - N_k(t-1)$ is the observed change in employment in occupation k. The approach requires observing only JtJ flows and exits to non-employment; entries are pinned down by the identity.

Because the method recovers entries as a residual, it does not require classifying pre-entry NE spells. Workers who have not yet entered the panel (panel-entry year after t) never contribute to JtJ or exit flows, regardless of how their pre-entry status is coded.[10]

[10]A fully panel-based alternative (Appendix C.2) directly tabulates the flow matrix and reads the non-

The flow matrix F is computed separately for each consecutive year-pair. To aggregate across years, we use *simple-average* conditional matrices: the annual conditional (row-normalised) matrices are averaged across year-pairs, giving equal weight to each year.

We now decompose the observed changes in employment shares $E_k(t)$. The change in k's employment share, $\Delta E_k(t) \equiv N_k(t)/N_e(t) - N_k(t-1)/N_e(t-1)$, can be rewritten as

$$\Delta E_k(t) \;=\; \frac{\Delta N_k(t) - E_k(t)\cdot \Delta N_e(t)}{N_e(t-1)}.$$

The numerator measures whether occupation k has grown faster or slower than the aggregate: $\Delta E_k > 0$ if and only if $\Delta N_k(t) > E_k(t)\cdot\Delta N_e(t)$, i.e. occupation k's growth exceeds the level that would keep its current share unchanged. The denominator $N_e(t-1)$ converts the difference into share-point units.

By the accounting identity (5), the net inflow $\Delta N_k(t)$ separates into the job-to-job component $\text{Net\_JtJ}_k(t)$ of equation (4) and a non-employment component capturing the net flow between k and non-employment (entries minus exits),

$$\text{Net\_NE}_k(t) \;\equiv\; F_{o,k}(t) - F_{k,o}(t). \tag{7}$$

Substituting these definitions into the share-change formula yields

$$\Delta E_k(t) \;=\; \frac{\text{Net\_JtJ}_k(t) + \text{Net\_NE}_k(t) - E_k(t)\cdot\Delta N_e(t)}{N_e(t-1)}. \tag{8}$$

In what follows, we focus on whether the net flow between occupation k and non-employment exceeds or falls short of the absorption that would maintain k's share given aggregate employment growth. We define the *excess* non-employment net flow as

$$\text{Excess\_NE}_k(t) \;\equiv\; \text{Net\_NE}_k(t) - E_k(t)\cdot\Delta N_e(t). \tag{9}$$

We also decompose the *Net_JtJ* component into bilateral net flows with each of the other employed occupations.

4 Employment Polarization in France

This section presents our main accounting results. We begin by tracking the evolution of the employment shares and document a clear break in 1994. We then use the framework of Section 3 to decompose worker flows across occupations and in and out of the labor

employment stock directly from the matched panel, under an explicit coding of pre-entry non-employment, rather than recovering the entries as a residual and using the INSEE rate. It returns identical job-to-job flows, exits, entries, and ΔE_k on all one-year pairs (the single two-year 1989 $\rightarrow$ 1991 gap pair differs negligibly, with ΔE_k unchanged; Appendix C.2); the only difference is the non-employment-persistence ($o \rightarrow o$) cell, which the share decomposition never uses.

market, and we show that the net flows from non-employment, rather than the job-to-job reallocation flows commonly emphasized in the literature, play the central role in the polarization that emerges after 1994. We next disaggregate routine occupations into their cognitive and manual sub-groups and show that the two behave markedly differently across the 1994 break. Finally, we offer what we view as the most plausible trigger of French polarization — the price of ICT equipment and services — whose timing aligns closely with the labor-market patterns we document.

4.1 Employment Shares and the 1994 Break

This subsection reports the aggregate employment-share evidence, both the within-employment shares and the augmented shares, and asks when France exhibits the polarization pattern documented in the literature.

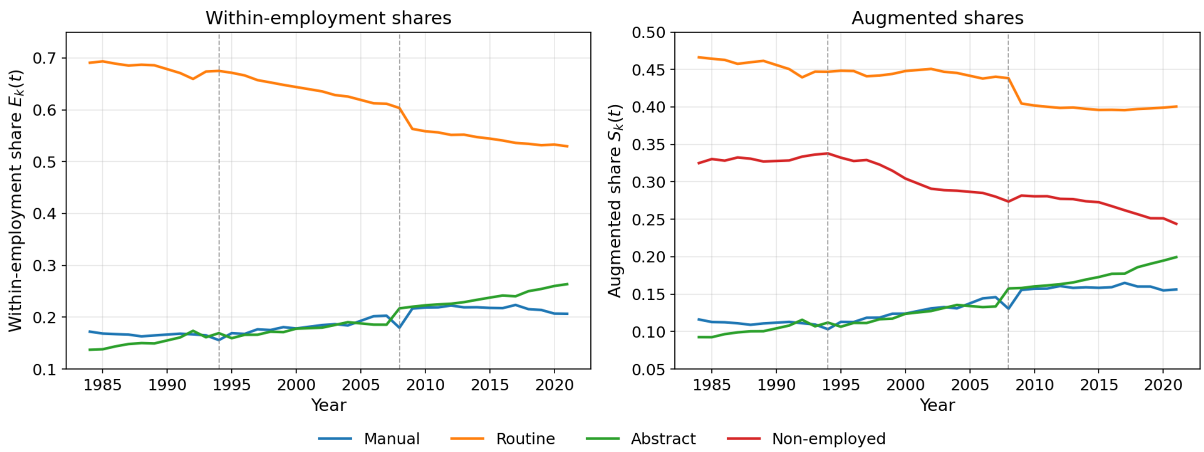

Figure 2: Occupational employment shares, workers aged 25–64. *Left:* within-employment shares $E_k(t)$. *Right:* augmented shares $S_k(t) = E_k(t) \cdot e(t)$ together with the non-employed share $S_o(t) = 1 - e(t)$, using the INSEE 25–64 employment rate. Dashed vertical lines mark 1994 and 2008.

Employment Shares. The left panel of Figure 2 plots the within-employment shares $E_k(t)$ over 1984–2021 for workers aged 25–64. Over the full sample, the routine share falls from 69.0% to 53.0%, the abstract share rises sharply from 13.7% to 26.4%, and the manual share rises from 17.2% to 20.7%. The sub-period numbers sharpen the timing. The routine share barely moves before 1994 (69.0% in 1984, 67.4% in 1993), then drops to 56.3% in 2009 and continues a more gradual decline to 53.0% in 2021: the bulk of the contraction occurs after 1994 and decelerates following the Great Recession. The abstract share rises from 13.7% in 1984 to 16.3% in 1993, accelerates to 22.0% in 2009, and reaches 26.4% in 2021. The manual share traces a U-shape across the break: slightly declining through the early 1990s (17.2% in 1984, 16.5% in 1993), reversing from 1994 onward, reaching 21.7% at the 2009 recession trough, and settling at 20.7% in 2021. All three series change regime together around 1994.

The right panel of Figure 2 plots the augmented shares $S_k(t)$, which bring out two features that the within-employment series hide. First, the non-employed share rises from 32.5% in 1984 to 33.6% in 1993 (reflecting the early-1990s French recession), then falls steadily to 24.4% by 2021, driven by the sustained rise in female labor-force participation. Second, the augmented routine share declines only moderately, from 46.6% in 1984 to 43.8% in 2008 (about 2.8 pp over 24 years), drops by a further 4 pp during and immediately after the 2008–2009 crisis to 39.6% in 2015, and remains roughly flat around 40% through 2021. What reads as a 16 pp fall in the within-employment routine share E_r is, in augmented terms, a 6 pp erosion in S_r concentrated in the crisis years.[11] The rise of abstract and the fall of non-employment are both visible on this scale, suggesting that a meaningful share of the reallocation runs through the non-employment margin.

Table 1 reports the simple-average annual change ΔE_k over 1984–1993 and 1994–2021 separately.

Table 1: Mean annual ΔE_k by sub-period (pp/year, simple-average)

Period	ΔE_m	ΔE_r	ΔE_a
Pre-polarization 1984–1993	−0.20	−0.09	+0.29
Polarization 1994–2021	+0.12	−0.52	+0.40
Full sample 1984–2021	+0.03	−0.41	+0.37

Notes: Values ×100 (pp/year). The Net_JtJ and Excess_NE decomposition of the same changes is presented in Section 4.2 (Table 2).

The share changes look different across the two periods on every dimension. Over 1984–1993, the routine share falls slightly ($\Delta E_r = -0.09$ pp/year), the manual share declines moderately ($\Delta E_m = -0.2$), and the abstract share rises at a moderate pace ($\Delta E_a = +0.29$). The 1984–1993 pattern is closer to a one-sided shift toward the high end of the occupational distribution than to polarization proper: manual is falling rather than rising, routine is flat, and only abstract grows. The U-shape that defines polarization — both tails rising while the middle contracts — emerges only from 1994. We therefore label the period "pre-polarization". From 1994 onward, all three shares change behavior substantially: the routine share contracts roughly five times faster ($\Delta E_r = -0.52$ pp/year), the manual share switches sign to positive ($\Delta E_m = +0.12$), and the abstract share accelerates to $\Delta E_a = +0.40$.

[11] The crisis-concentrated decline with no subsequent recovery is consistent with the jobless-recoveries pattern documented for the United States by Jaimovich and Siu (2020), who show that routine employment losses are concentrated in recessions and are not reabsorbed afterward.

4.2 Flow Decomposition and the Anatomy of Polarization

We now use the flow accounting identities of Section 3 to decompose the changes in the employment shares into a job-to-job component (Net_JtJ) and a non-employment component (Excess_NE). The aim is to weigh the job-to-job reallocation flows commonly emphasized in the literature against the entry and exit margins to and from the labor market in driving the polarization observed since 1994.

Full-period decomposition. Figure 3 reports the bilateral Net_JtJ decomposition over the full 1984–2021 sample, with the Excess_NE channel added and the resulting ΔE_k marked.

The routine decline ($\Delta E_r = -0.41$ pp/year) is carried by the Net_JtJ channel (-0.41 pp/year, see Table 2). The bilateral $r \to a$ flow (-0.53 pp/year) alone is larger in absolute terms than the entire Net_JtJ loss, with a positive net inflow from manual ($+0.11$ pp/year) partly offsetting the upgrading drain.

Conversely, the relatively small manual rise ($\Delta E_m = +0.03$ pp/year) is the results of two almost offsetting net flows: a positive Excess_NE channel ($+0.19$ pp/year) and a negative Net_JtJ channel (-0.16 pp/year): non-employed workers enter manual occupations at rates above the manual employment share, while manual workers are net losers from direct switches.

Finally, the abstract rise ($\Delta E_a = +0.37$ pp/year) is carried almost entirely by Net_JtJ ($+0.57$ pp/year, nearly all from routine), with a small negative Excess_NE (-0.19) reflecting that exits to non-employment are disproportionately from abstract workers. The full-period decomposition by occupation and channel, together with the bilateral Net_JtJ split, is reported in Appendix B.2.

The 1994 break. Table 2 splits the full-period decomposition into the pre-polarization (1984–93) and polarization (1994–2021) sub-periods. Figure 4 displays the same split graphically.

We highlight three main features of Table 2. First, the $r \to a$ upgrading drain is essentially stable across the break: net $r \to a$ runs at -0.54 pp/year pre-polarization and -0.52 pp/year during polarization, essentially stable. The slow, steady drain of routine workers into abstract is a structural feature of the French labor market; it is not what switches on in 1994. Second, the manual-to-routine Net_JtJ buffer collapses across the break: net $m \to r$ falls from $+0.20$ pre-polarization to $+0.08$ pp/year during polarization. The aggregate routine Net_JtJ channel becomes more negative (-0.34 vs. -0.44) because the constant $r \to a$ drain is less offset by the weakened $m \to r$ buffer. Third, the routine Excess_NE channel flips sign, from $+0.25$ pre-polarization to -0.09 during polarization. Before 1994, the non-employment channel buffers routine, partially offsetting the Net_JtJ

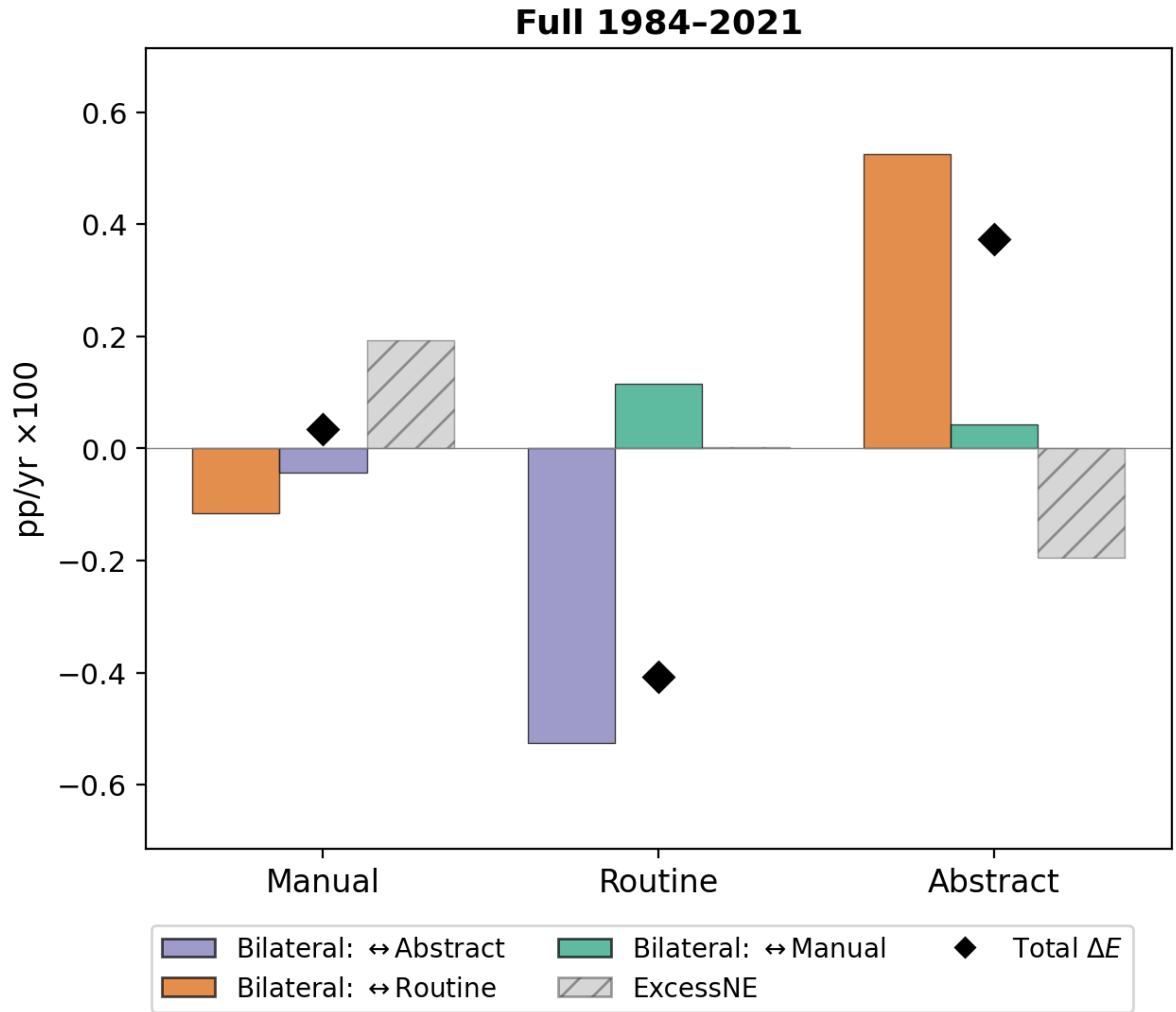

Figure 3: Mean annual bilateral Net_JtJ decomposition together with the Excess_NE channel and the resulting ΔE_k, full period 1984–2021. For each destination occupation k, the three solid bars show the bilateral net flows from the other two employed states; the hatched bar is Excess_NE; the diamond is ΔE_k.

Table 2: Mean annual bilateral decomposition by sub-period (pp/year)

Period		Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
Full sample 1984–2021	manual	—	−0.11	−0.04	−0.16	+0.19	+0.03
	routine	+0.11	—	−0.53	−0.41	+0.00	−0.41
	abstract	+0.04	+0.53	—	+0.57	−0.19	+0.37
Pre-polarization 1984–1993	manual	—	−0.20	−0.03	−0.24	+0.04	−0.20
	routine	+0.20	—	−0.54	−0.34	+0.25	−0.09
	abstract	+0.03	+0.54	—	+0.58	−0.29	+0.29
Polarization 1994–2021	manual	—	−0.08	−0.05	−0.13	+0.25	+0.12
	routine	+0.08	—	−0.52	−0.44	−0.09	−0.52
	abstract	+0.05	+0.52	—	+0.57	−0.16	+0.40

Notes: Values ×100 (pp/year). "Net-k'" reports the bilateral net inflow $(F_{k',k} - F_{k,k'})/N_e(t-1)$; Net_JtJ is the sum of the three bilateral entries. Each panel sums to zero across occupations by construction.

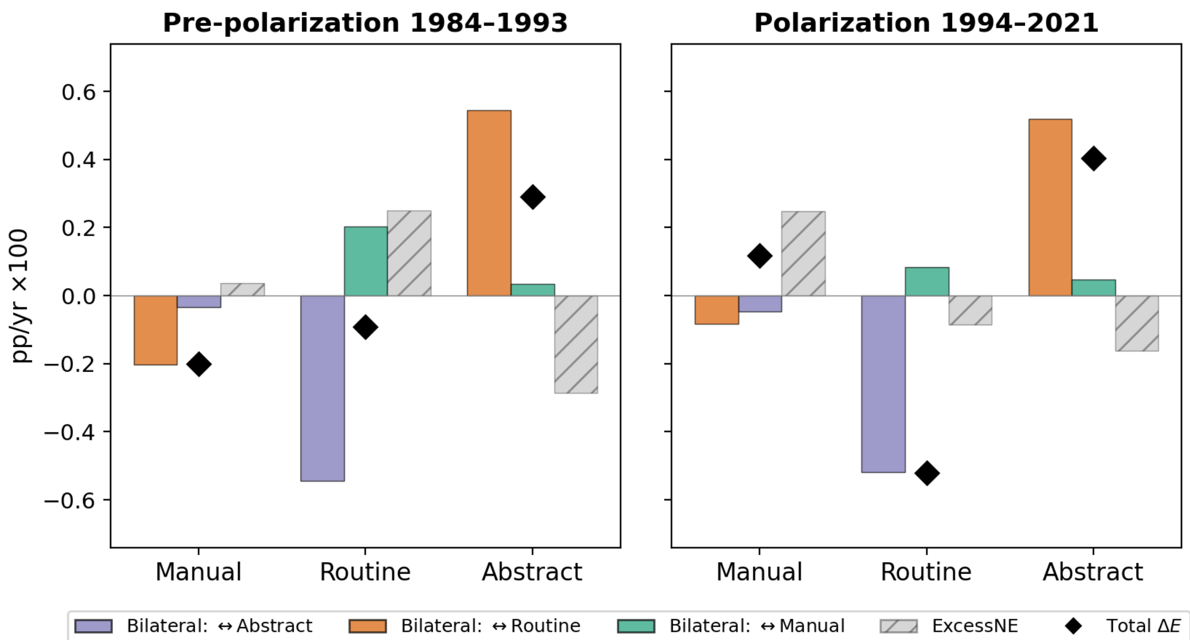

Figure 4: Mean annual bilateral decomposition: pre-polarization (1984–1993) vs. polarization (1994–2021).

drain; after 1994, the buffer disappears. The sign flip accounts for most of the acceleration in the routine decline (from −0.09 to −0.52 pp/year).

The polarization regime itself is not homogeneous over time. The routine decline is steepest during the 2007–09 crisis, which pulls the full-period Excess_NE level to −0.09, more negative than in the years on either side. A more robust statement than a level "sign flip," whose exact crossing of zero is sensitive to those crisis years, is that the non-employment buffer simply *disappears* around 1994 and stays absent thereafter— the crisis amplifying the magnitude of the decline rather than its composition (Appendix B.2). The qualitative pattern — stable $r \to a$ drain, collapse of both the $m \to r$ and Excess_NE buffers around 1994 — is robust to the age window (Appendix C.1).

4.3 The routine Excess Non-employment collapse: entry or exit?

A natural question is whether the routine Excess_NE sign flip reflects a change on the entry margin (fewer workers arriving in routine from non-employment), on the exit margin (more workers leaving routine to non-employment), or both.

To separate the two, measure routine's non-employment flows against their share-neutral benchmarks $E_r \sum_k F_{o,k}$ and $E_r \sum_k F_{k,o}$, i.e. respectively those entry and exit flows from and to non-employment that would keep the routine employment share constant. Defining $\rho_r^{\text{ent}} = F_{o,r}/(E_r \sum_k F_{o,k})$ and $\rho_r^{\text{exit}} = F_{r,o}/(E_r \sum_k F_{k,o})$ routine's inflow and outflow as fractions of those benchmarks, the routine Excess_NE channel—its contribution to ΔE_r—can be written as

$$\frac{\text{Excess\_NE}_r(t)}{N_e(t{-}1)} = \underbrace{\frac{E_r(t)\sum_k F_{o,k}(t)}{N_e(t{-}1)}}_{\text{benchmark entry}} \underbrace{(\rho_r^{\text{ent}} - 1)}_{\text{deviation entry}} - \underbrace{\frac{E_r(t)\sum_k F_{k,o}(t)}{N_e(t{-}1)}}_{\text{benchmark exit}} \underbrace{(\rho_r^{\text{exit}} - 1)}_{\text{deviation exit}}.$$

We aim at decomposing the −0.34 reduction in this channel across 1994 in the entry and exit contribution. To this purpose, the change of both the entry and the exit components in the above expression can be decomposed as:

$$\Delta[\text{benchmark} \cdot (\rho - 1)] = \underbrace{(\bar{\rho} - 1)\,\Delta\text{benchmark}}_{\text{benchmark effect}} + \underbrace{\overline{\text{benchmark}}\,\Delta(\rho - 1)}_{\text{deviation effect}},$$

with bars denoting averages between adjacent periods. Table 3 reports this split for the entry margin, the exit margin, and their difference (Net = Entry − Exit), the last being the channel's own −0.34 swing.

The benchmark effect nearly cancels in the net (−0.04). The deviation $\bar{\rho} - 1$ is negative (the $\rho_r - 1$ column), and the benchmark $E_r \sum_k F$ falls — routine's share E_r drops from 69% to 53% and the aggregate entry and exit rates fall (16.0% → 12.1% and 13.9% → 10.3% of employment) — so the product is positive: +0.16 on entry, +0.20 on exit. The channel

Table 3: Benchmark and deviation entry and exit

Routine margin	Deviation ρ_r-1 (pre→pol)	Deviation effect	Benchmark effect	Δ
Entry	$-0.03 \to -0.06$	-0.31	$+0.16$	-0.15
Exit	$-0.06 \to -0.06$	-0.01	$+0.20$	$+0.19$
Net (Excess_NE)	—	-0.30	-0.04	-0.34

nets entry against exit, so these two close effects largely cancel: $+0.16 - 0.20 = -0.04$.

The remaining -0.30 pp/year is the deviation effect, and it is essentially all entry: routine's entry deviation widens across 1994, $\rho_r^{\text{ent}} - 1 = -0.03 \to -0.06$, while its exit deviation is flat, $\rho_r^{\text{exit}} - 1 = -0.06$ throughout. The routine sign flip is therefore a change in routine's deviation on the entry margin—a shrinking fraction of the non-employed flow into routine relative to its share-neutral benchmark—not a faster release of routine workers to non-employment, whose exit deviation relative to the benchmark does not move.

A counterfactual isolates the two deviations. In the polarization period routine declines at $\Delta E_r = -0.52$ pp/yr. Had its exit deviation held at its pre-1994 value—benchmarks and entry deviation moving as observed—the decline would be essentially unchanged, -0.53. Had its entry deviation held instead, the decline would more than halve, to -0.21: with routine no longer increasingly passed over in non-employment hiring, the channel would have stayed a buffer rather than turning to a drag. The non-employment component of the routine decline is, by this measure, an entry-deviation phenomenon.

Direct vs. indirect downgrading. A natural concern about reading the Excess_NE channel as a clean measure of non-employment flows is that it may hide $r \to m$ downgrading that runs through a non-employment spell: a routine worker who loses her job, spends one period in non-employment, and re-enters as manual would appear in the accounts as an $r \to o$ exit followed by an $o \to m$ entry, not as a direct $r \to m$ downgrade. In Appendix B.2.3 we address this concern by constructing two-step transitions via non-employment, e–o–e. The via-o flows are *net upgrading* on every bilateral pair (m–r, m–a, and r–a), and the pattern intensifies during polarization. The direct $r \to m$ displacement channel that a routine-displacement account would predict is therefore absent whether one looks at one-step Net_JtJ flows or at indirect paths through non-employment.

In sum, the $r \to a$ upgrading drain is essentially stable across 1994. Polarization in France from 1994 is driven primarily by the collapse of the non-employment-to-routine entry rate, which flips the sign of routine's Excess_NE from positive to negative. The accompanying collapse of the manual-to-routine upgrading flow contributes as well, but plays a smaller role.

4.4 Asymmetries Inside Routine Occupations

To dig further into the behavior of routine employment, we now separate routine occupations into two sub-groups, *routine cognitive* (rc) and *routine manual* (rm). The former comprises white-collar occupations; the latter, blue-collar.[12]

Figure 5 plots the within-employment shares of the five sub-groups. The rm share falls continuously from 34.3% of employment in 1984 to 18.3% in 2021, at a rate broadly consistent with the secular contraction of French manufacturing employment. The rc share follows a different path: it *rises* from 34.8% in 1984 to 42.0% in 1994 (about +0.72 pp/year), then turns and drifts back down to 34.7% by 2021 (about −0.27 pp/year). The aggregate routine plateau before 1994 therefore reflects rc growth roughly canceling rm contraction. The aggregate routine decline after 1994 reflects rc switching from growth to decline while rm continues to contract, so that the two sub-groups stop offsetting each other and start moving in the same direction. In this sense the 1994 break in routine is located on the rc side: rm was already declining and continues to do so; what changes in 1994 is the rc trajectory.

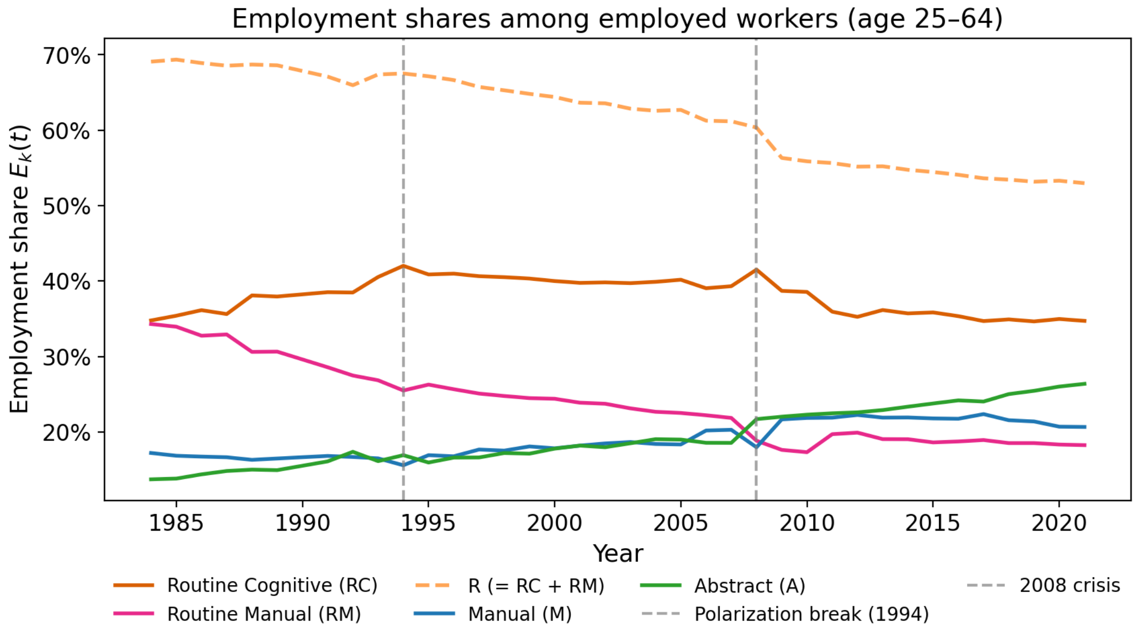

Figure 5: Within-employment shares, five-state system. The aggregate routine share (dashed) is the sum of rc (routine cognitive) and rm (routine manual).

[12]We use the CS2 classification. *Routine cognitive* (rc) comprises the *professions intermédiaires* and *employés qualifiés* (CS2 $\in \{42, 43, 45, 46, 48, 52, 54\}$), i.e., white-collar clerical, administrative, and supervisory occupations. *Routine manual* (rm) comprises *ouvriers qualifiés* and *non qualifiés* of industrial and artisanal type (CS2 $\in \{62, 63, 65, 67\}$), i.e., blue-collar production workers. The two sub-groups rank in the wage distribution as $rc > rm$: rc wages run at 1.5–1.7 times the manual average and rm wages between 1.2 and 1.5 times the manual average (Appendix A.1).

4.4.1 Flows by Routine Sub-group

We now extend the flow decomposition from three to four employed states, treating the two routine sub-groups separately. Figure 6 reports the resulting five-state decomposition.[13]

Looking across the two panels, three features are worth noting. First, the routine Excess_NE sign flip is driven by routine-cognitive occupations and partially mitigated by routine-manual ones. For rc, Excess_NE collapses from +0.80 pp/year pre-1994 to 0.04 after 1994, a swing of −0.76 pp/year, visible in Figure 6 as the disappearance of the tall hatched bar in the rc column between the two panels. For rm, Excess_NE was already strongly negative before 1994 (−0.56), and the negative pull weakens to −0.12 after 1994. The positive rm swing of +0.44 pp/year partially offsets the routine fall driven by the rc collapse.

Decomposing the entry residual of Section 4.2 by sub-group confirms the same reading: rc's excess entry collapses by −0.53 pp/year (from +0.28 pre-polarization to −0.25 during polarization), while rm's excess entry increases by +0.45 pp/year, as rm's employment share falls faster than its non-employment entry rate. The entry-side collapse of the aggregate routine Excess_NE channel is therefore entirely an rc phenomenon.

Before 1994, non-employed workers consistently fed rc occupations — educated new entrants starting careers in clerical, administrative, and supervisory jobs, often as a stepping stone toward abstract — while non-employment consistently drained rm occupations, as France's deindustrialization eroded demand for production workers. What changes in 1994 is that rc stops compensating for the continued drain in rm. As the flow of entries from non-employment into rc collapses, rc stops expanding; rm contraction attenuates somewhat, but the combined routine stock $rc + rm$ enters its sustained decline — the polarization phase.

Second, the $r \to a$ upgrading drain runs mostly through rc. The purple $\leftrightarrow a$ bars in the rc and rm columns of Figure 6 differ sharply: net $rc \to a$ is +0.43 pp/year pre-polarization and +0.40 during polarization, while net $rm \to a$ is much smaller, +0.12 in both sub-periods. Roughly three-quarters of the aggregate $r \to a$ upgrading drain in each sub-period runs through the cognitive side of routine, consistent with the CS2-level concentration in Appendix B.3, whose dominant fine-grained pair is CS46→CS37 (intermediate administrative professions upgrading to cadre), an $rc \to a$ transition.

Third, the manual-to-routine Net_JtJ buffer is also an $m \to rc$ flow, not $m \to rm$ as one might expect. In both panels the green $\leftrightarrow m$ bar is substantially larger for rc than for rm: net $m \to rc$ is +0.20 pp/year pre-polarization and falls to +0.08 during polarization, whereas net $m \to rm$ is much smaller (+0.01 pre-polarization and close to zero after). The collapse of the aggregate $m \to r$ buffer documented in Section 4.2 is therefore primarily an $m \to rc$ phenomenon, not a weakening of the blue-collar ladder from manual service

[13]The conditional and unconditional transition matrices for the 5×5 system across sub-periods are reported in Appendix B.3.

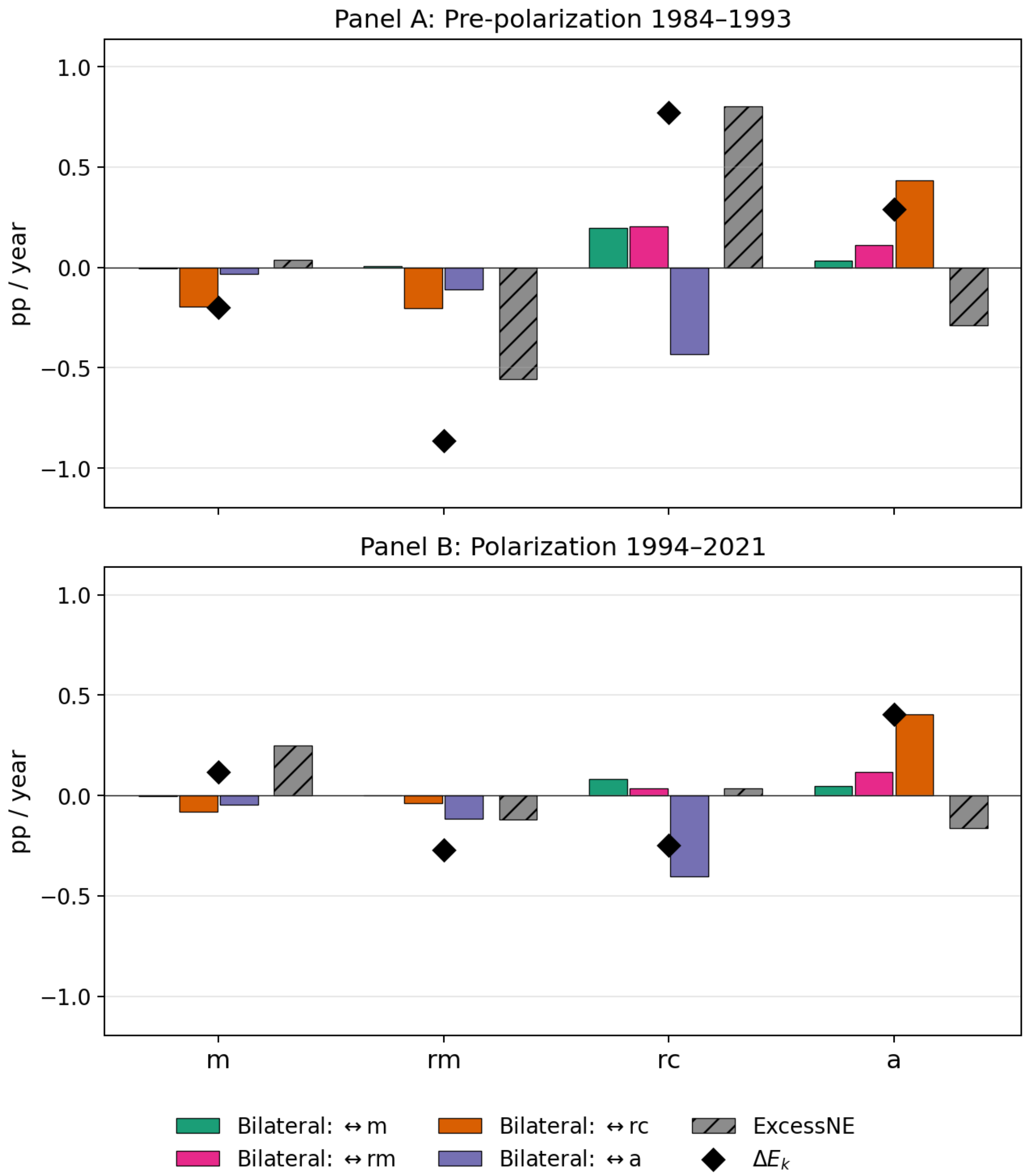

Figure 6: Mean annual bilateral decomposition of ΔE_k for the five-state system (m, rm, rc, a). Panel A: pre-polarization 1984–1993. Panel B: polarization 1994–2021. For each destination occupation, colored bars show the bilateral Net_JtJ flows from the other three employed states (green: $\leftrightarrow m$; pink: $\leftrightarrow rm$; orange: $\leftrightarrow rc$; purple: $\leftrightarrow a$); the hatched bar is Excess_NE; the diamond marks ΔE_k. Destinations are ordered along the wage hierarchy $m < rm < rc < a$. in the initial year.

work into manual routine production. The CS2 decomposition in Appendix B.3 clarifies the mechanism: the dominant fine-grained $m \to r$ pair is CS55→CS46 (*employés de commerce* upgrading to *professions intermédiaires*), a service-to-white-collar transition that runs entirely through the cognitive side of routine rather than the manual side. The 1994 break therefore weakens not only the *rc* upgrading ladder into abstract but also the service-worker upgrading ladder into *rc* itself.

4.5 A candidate trigger: ICT prices

This paper is deliberately agnostic on the causes of labor market polarization in France. However, we find appropriate to discuss some potential determinants of it. A natural class of explanations for an *rc*-specific break around 1994 involves the cost of the technologies that substitute most directly for cognitive routine tasks: office computing on the one hand, and telecommunications on the other [14]. Figure 7 plots the INSEE consumer price indices for these two categories, both base 2015 = 100, from 1990 to 2023. Panel A reports COICOP 08.3, communications services (telephone and telefax, postal services). Panel B reports COICOP 09.1.3, information processing equipment (desktop and portable computers, peripherals, pre-packaged software), on a logarithmic scale.

The two series carry different information. Communications prices were held roughly constant through the early 1990s by the regulated-monopoly structure of the French telecoms sector, and then fell sharply once France opened the market to competition (EU directives of 1996 and 1998, France Télécom privatization in 1997, mobile market liberalization). The peak of Panel A occurs in 1994, and the sustained decline begins the year after. Computer equipment, by contrast, had been cheapening continuously for two decades by the time our data begin in 1990, but an acceleration of the log-linear decline is visible after 1994 (the annual rate of decline roughly doubles between 1990–94 and 1994–2000). Both series therefore exhibit a change in trajectory around 1994, coinciding with the labor-market break documented above.

The timing and sub-group specificity of the *rc* break documented in Section 4.4 are

[14] A France-specific fiscal reform also coincides closely with the 1994 break. From July 1993 the government introduced general employer social-contribution relief on low wages: family-allowance contributions were exempted below 1.1 times the minimum wage (SMIC) (5.4 percentage points of employer contributions) and halved between 1.1 and 1.2 times the SMIC (2.7 percentage points), a schedule progressively extended up the wage distribution by subsequent reforms, reaching 1.6 times the SMIC today. By cheapening low-wage relative to middle-wage labor, this reform is a plausible demand-side contributor to the post-1994 rise in non-employment-to-manual entry documented in Section 5. It is unlikely, however, to drive the whole polarization pattern: the evaluation literature finds the employment effects of these schemes substantial for low-wage labor and much more mixed higher up the distribution (Bozio and Wasmer, 2024). The reform should therefore be seen as complementary to the ICT-price trigger rather than competing with it — the two act on different margins, the ICT price on the routine decline and the low-wage relief on the manual rise, though by steering non-employment entrants toward cheaper manual jobs the relief could also feed the routine entry collapse. Disentangling the two — for instance by testing whether the entry collapse tracks the discrete schedule expansions or the smooth ICT-price trend — is beyond the scope of the paper and is left to future work.

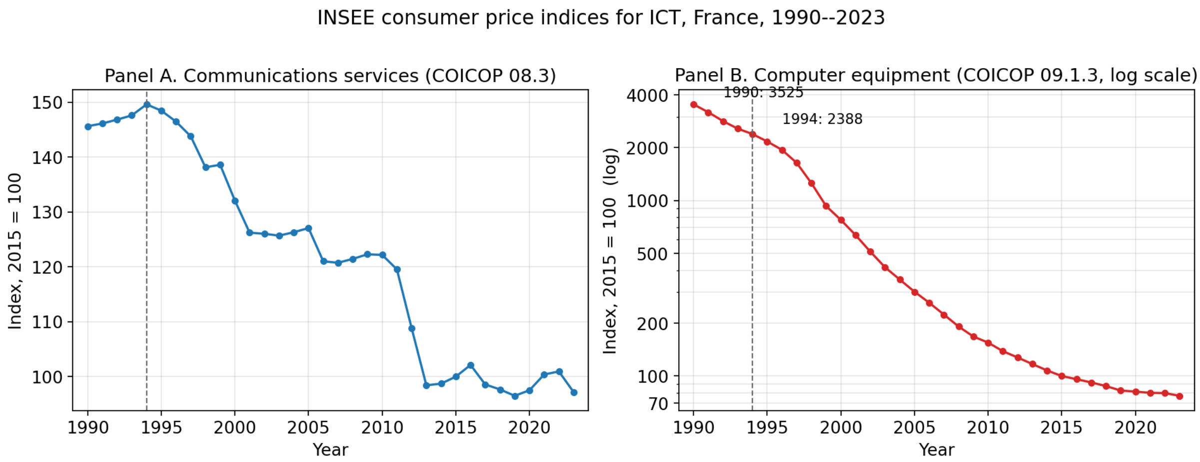

Figure 7: Consumer price indices for ICT, France, 1990–2023 (INSEE IPC, base 2015 = 100). Panel A: communications services (COICOP 08.3, idBank 001764308). Panel B: information processing equipment (COICOP 09.1.3, idBank 001764157), log scale.

suggestive in this context. Cognitive routine tasks — spreadsheets, word processing, databases, accounting systems — are more directly substitutable by the ICT bundle (computing plus telecoms) than production tasks, and the 1994 telecoms-price break appears on the same side of the labor-market decomposition as the flow anomaly (rc rather than rm).

5 Heterogeneity: Education, Gender, Tenure

In this section we examine whether polarization affected different groups of workers differently. We explore how heterogeneity in education and other covariates drives differences in job flows. First, we stratify the aggregate ΔE_k decomposition by college status. The salient result is that the rise of polarization is largely a non-college phenomenon. We then examine the partial correlations of college, female, and tenure with one-year transitions across occupations and labor-market participation. We find notable differences in how gender and tenure shape the impact of polarization across workers.

5.1 Education

We first ask whether college and non-college workers behave differently. Such differences are central to dominant views of labor-market polarization. Under Autor and Dorn (2013) (AD2013 hereafter), routine-biased technological change displaces routine workers, and the two education groups follow different trajectories: college-educated workers and college-educated (re-)entrants move more strongly into abstract occupations, while non-college workers (incumbents and entrants) move more into manual service occupations or spend longer spells in non-employment. If so, the rise of abstract employment would be largely a college phenomenon. We test whether the French experience follows these patterns. This exercise is also related to Cortes et al. (2017), who show for the United States that the decline of routine employment is concentrated among specific demographic groups that also account for much of the rise in non-employment and low-wage non-routine manual work.

Education is recorded in DADS only from 2002 onward. For earlier years we project the first observed education level backward via panel identifiers, which mechanically restricts the pre-polarization, education-stratified sample to workers who survive to 2002. The resulting college share (among employed workers aged 25–64) rises from 10.9% in 2002–05 to 15.0% in 2017–21. About 75% of unique workers in the panel have an education observation, covering about 91% of all Net_JtJ observations because survivors have longer panel histories than non-survivors. Despite these limitations, the recovered patterns illustrate the behavioral differences between college and non-college workers.

Figure 8 plots the within-employed occupational distribution by education group. A gray reference line shows the within-employed shares computed on the full age-25–64 sample (regardless of whether the worker's education is observed); the gray line tracks the black all-with-education line closely throughout the panel, indicating that the aggregate occupational composition is not materially biased by the educ-observation selection.[15]

Within each education group, the patterns differ substantially. Among non-college workers the typical U-shape of polarization emerges: between 1984 and 2021 the routine share falls by 13.3 pp (from 73.0% to 59.6%) while both the manual share (+7.7 pp, from 16.7% to 24.4%) and the abstract share (+5.7 pp, from 10.4% to 16.0%) rise. Among college workers, by contrast, the within-employed occupational distribution is considerably flatter. No clear polarization pattern is visible in the time series; if any, the data show only a mild upgrading, concentrated in the pre-1994 period. Between 1984 and 2021 the abstract share rises by 7.0 pp (from 58.5% to 65.5%), the routine share declines by 6.4 pp (from 37.2% to 30.8%), and the manual share is essentially flat (−0.6 pp, from 4.2% to 3.7%).

An Oaxaca–Blinder decomposition of each aggregate share change between 1984 and 2021 (within the educ-observed sample) quantifies the relative weight of within-group reallocation versus the rising college share (Appendix B.2.4; Table 16 reports the full breakdown). The aggregate abstract rise of +11.5 pp splits roughly equally between within-non-college upgrading (+5.0 pp, 43% of the total) and the rising college share (+5.7 pp, roughly 50%), with a small contribution from within-college upgrading (+0.8 pp, 7%). The

[15]Educational attainment is not directly recorded in the data for the pre-2002 portion of the panel; for those years, education is inferred by backward propagation from each worker's first post-2002 observed value. This restricts the pre-2002 by-education sample to workers who remain in the panel until at least 2002. Even after 2002, a non-trivial share of employed workers has no observed educational attainment at any point, rising from 4% in 2002 to 18% by 2021. Appendix C.5 reports bounded estimates for non-college occupational shares under the maximal-imputation assumption that all workers with no observed educational attainment are non-college.

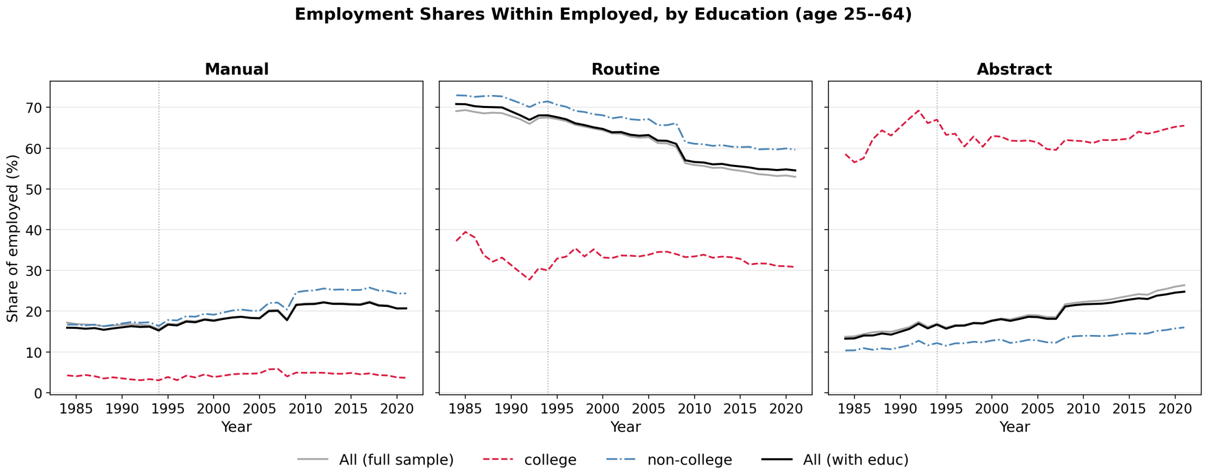

Figure 8: Within-employed occupational shares by education group. Each panel plots manual, routine, and abstract shares among employed college and non-college workers over 1984–2021.

aggregate routine decline (−16.3 pp) is overwhelmingly a within-non-college phenomenon (−11.8 pp, 72%), with a smaller compositional contribution from the rising college share (−3.8 pp, 23%) and a marginal within-college contribution (−0.8 pp, 5%). The aggregate manual rise of +4.8 pp reflects a sizable within-non-college increase (+6.8 pp), partly offset by the compositional shift toward the (lower-manual) college pool (−1.9 pp); the within-college trajectory is essentially flat.[16]

Figures 9 and 10 report the simple-average ΔE_k decomposition by education group, with pre-polarization (1984–1993) and polarization (1994–2021) panels placed side by side within each figure and sharing the y-axis. The full bilateral Net_JtJ flows underlying the two figures are reported in Appendix B.2.4.

The most striking pattern is the steadiness of the $r \to a$ flow within the non-college sample. The bilateral net flow from routine to abstract for non-college workers is +0.47 pp/year in 1984–93 and +0.45 pp/year in 1994–2021—essentially unchanged across the two regimes. The routine-to-abstract upgrading channel therefore operates among non-college workers throughout the panel rather than emerging with polarization. Inside the college sample the same flow runs much faster, at +2.17 pp/year pre-polarization and +0.98 pp/year during polarization. The college rate is roughly five times the non-college rate pre-1994 and about 2.2 times after; in either period, the college $r \to a$ premium is large. The non-college population is, however, large enough that its contribution to the aggregate abstract gain remains quantitatively important, consistent with the Oaxaca-Blinder de-

[16]A parallel literature reads the presence of non-college workers in high-skill occupations not as upgrading but as *vertical mismatch*. Garibaldi et al. (2025) classify non-college workers in majority-college occupations as "over-employed" and document that they account for a substantial share of non-college employment in the United States. Their occupations are ranked by the modal education of incumbents rather than by task content, so the two classifications need not coincide, and we take no stance on whether the rise in non-college abstract employment reflects mismatch or genuine task upgrading.

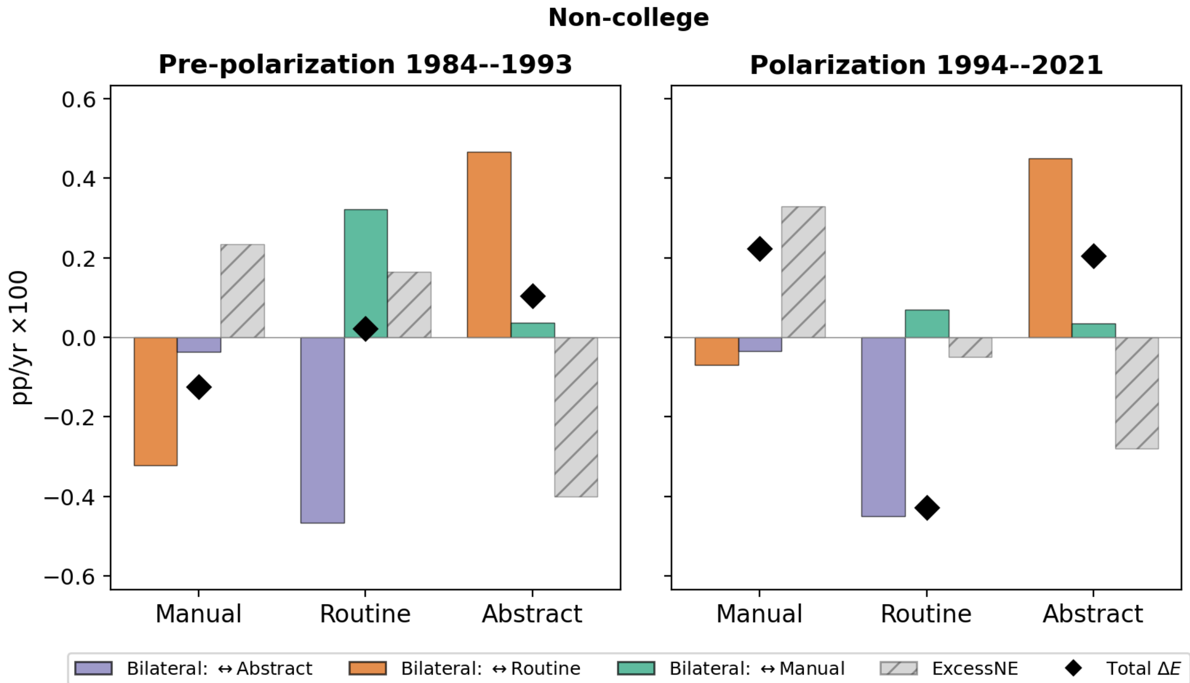

Figure 9: Mean annual decomposition for non-college workers, pre-polarization (left, 1984–1993) and polarization (right, 1994–2021, excluding DADS expansion pairs 2001–02 and 2005–06). Bars show bilateral Net_JtJ flows (net from m, r, a), the Excess_NE channel, and the resulting ΔE_k (diamond).

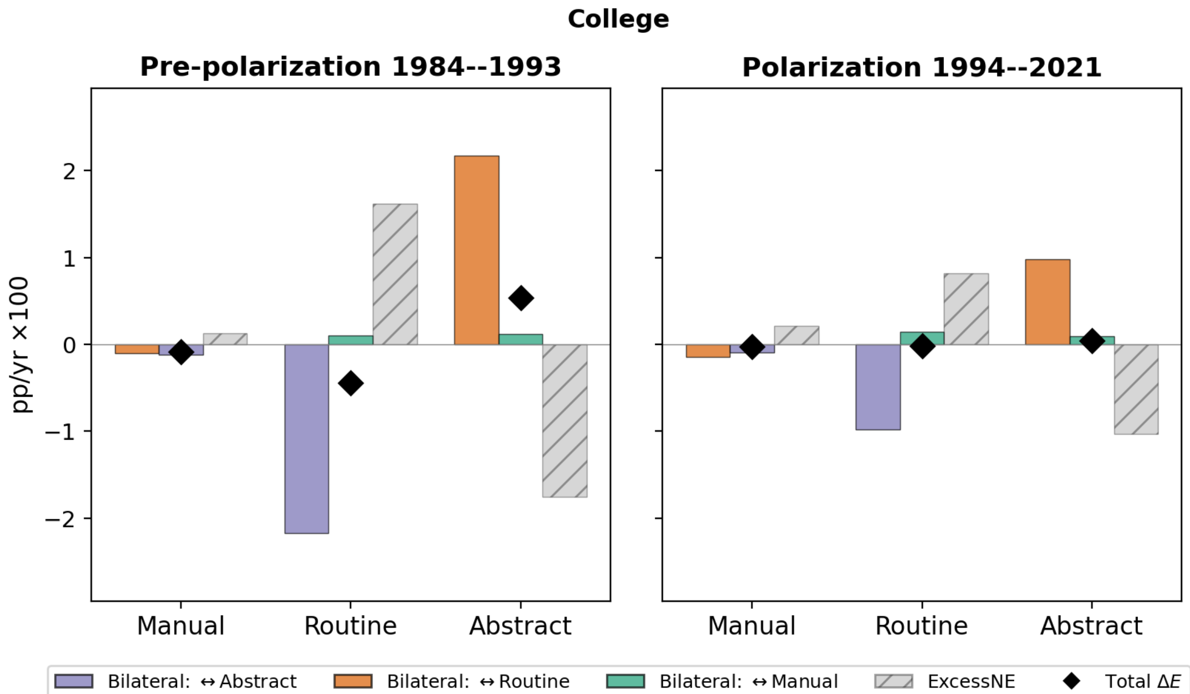

Figure 10: Mean annual decomposition for college workers.

composition above. Specifically, despite the lower per-worker transitional probability, non-college workers account for roughly 76% of all net routine-to-abstract worker-flows in both periods.

The cross-period change in ΔE_r inside the non-college sample, from +0.02 pp/year pre-polarization to −0.43 during polarization, does not come from a change in $r \to a$ upgrading but from two other margins. The manual-to-routine bilateral flow weakens markedly (+0.32 to +0.07 pp/year, a −0.25 swing) and the routine non-employment entry channel flips sign (+0.17 to −0.05 pp/year). The two together account for the full swing in non-college ΔE_r.

A symmetric reading of the manual share among non-college workers reinforces doubts about the displacement-into-manual view of routine decline, at least for France. The non-college ΔE_m rises from −0.13 pre-polarization to +0.22 pp/year during polarization, a positive swing of +0.35. The bilateral $r \to m$ flow inside the non-college sample is, however, negative in both periods (net $m \to r$ is +0.32 pp/year pre-polarization and +0.07 pp/year during polarization): routine workers are not displaced into manual on net. The post-1994 rise in non-college manual employment is instead driven by a sizeable non-employment entry channel into manual: non-college Excess_NE_m equals +0.33 pp/year, so the NE-entry channel is the entire net source of the post-1994 manual rise (+0.22 pp/year).

To sum up, the polarization break for non-college workers is a combination of a steady upgrading ladder, a weakening manual-to-routine buffer, and an entry-side reallocation away from routine, rather than a direct displacement of routine workers into manual employment.

All in all, the flow evidence supports a more optimistic reading of the routinization hypothesis of Autor and Dorn (2013): rather than non-college routine workers being either displaced into non-employment or downgraded into manual jobs, a persistent demand for abstract, ICT-complementary jobs, sustained a stable upgrading channel that continued to pull non-college routine workers up the occupational ladder at the same rate before and after 1994 – either because firms have consistently drawn on the non-college pool to fill abstract positions, or because technological progress has continuously induced routine workers to acquire abstract skills and move up the ladder. The two mechanisms are not mutually exclusive and properly disentangling their relative contributions requires quantitative models of dynamic occupational choice and human capital accumulation, disciplined by the kind of bilateral-flow and non-employment evidence we provide here.

5.2 Beyond Education: Gender and Tenure

The flow decomposition is occupation-level and does not identify which workers make each transition. We estimate linear probability models (LPMs) for one-year transitions from each origin state $\kappa \in \{m, r, a, o\}$ to each destination $\kappa' \neq \kappa$, isolating the partial association of education, gender, tenure, and age with each transition. Define binary transition indicators $y_{it}^{\kappa\to\kappa'} = 1$ if worker i in state κ at t is in state κ' at t+1. Our baseline specification is

$$y_{it}^{\kappa\to\kappa'} = \alpha + \beta_1\,\text{age}_c + \beta_2\,\text{female} + \beta_3\,\text{college} + \delta\,\text{tenure} + \text{Year FE} + \varepsilon_{it}, \qquad (10)$$

where age_c is age (entering linearly) and tenure is years in the current two-digit PCS code. For the three o-origin samples ($y^{o\to m}$, $y^{o\to r}$, $y^{o\to a}$) we restrict the sample to re-entrants from non-employment, i.e., workers with at least one prior observed year of employment in the panel.[17] As discussed in Section 3, workers whose first DADS appearance is directly in employment cannot be perfectly distinguished from workers who were already employed prior to being added to the sample. For these o-origin rows, tenure in equation (10) is replaced by the duration of the current observed non-employment spell ending at t (in years), with the standard re-employment hazard interpretation; for employed workers ($\kappa = m, r, a$), tenure must be interpreted as the number of years in the current CS2 employment category. The college indicator is a dummy variable taking value 1 when the worker has a college degree or higher, and zero otherwise; since education is not recorded before 2002, it is back-projected within worker from the first year it is observed as in the previous section. Every regression weights each year equally — the regression counterpart of the simple-average-over-pairs convention used throughout the paper.

5.2.1 Pooled Estimates

Table 4 reports the college, female, and tenure coefficients for the twelve one-year transitions, organized into four groups: within-employment upgrading, within-employment downgrading, exits to non-employment, and (re-)entries from non-employment. Full regression tables for all twelve transitions appear in Appendix B.4.

First, *education is the strongest predictor of upgrading.* The college coefficient on $r \to a$ is the largest in the table, +9.7 pp on an unconditional rate of 3.3%, so college roughly quadruples the annual routine-to-abstract rate. College also raises $m \to a$ upgrading (+7.7 pp over an unconditional rate of 1.1%) and re-entry into abstract ($o \to a$, +7.1 pp over 2.9%), while it lowers re-entry into routine and manual ($o \to r$, −4.7; $o \to m$, −4.2) and downgrading from abstract ($a \to r$, −6.5).

Second, *tenure is a stabilizer.* A one-year increase in occupation-specific tenure lowers the $r \to a$ rate by 0.31 pp, the $m \to r$ rate by 1.12 pp, and the $a \to r$ rate by 0.95 pp; longer non-employment spells lower re-entry on every destination. As workers accumulate occupation-specific human capital they move less, suggesting that polarization operates

[17] This keeps most of the entry flow: re-entrants account for about 63% of all non-employment-to-employment entries before 1994 and 77% after. The excluded categories are aging-in at age 25, around a quarter throughout, and panel first-entrants, whose share falls from about 10% to under 1%.

Table 4: LPM partial correlations by transition type: all twelve transitions

Transition	$\bar{y}$	**College**	**Female**	**Tenure**
Within-employment upgrading				
$m \to r$	0.084	$+0.057^{***}$	-0.038^{***}	-0.011^{***}
$r \to a$	0.033	$+0.097^{***}$	-0.022^{***}	-0.003^{***}
$m \to a$	0.011	$+0.077^{***}$	-0.009^{***}	-0.001^{***}
Within-employment downgrading				
$r \to m$	0.028	-0.015^{***}	-0.004^{***}	-0.004^{***}
$a \to r$	0.072	-0.065^{***}	$+0.036^{***}$	-0.009^{***}
$a \to m$	0.009	-0.008^{***}	$+0.002^{***}$	-0.001^{***}
Exits to non-employment				
$a \to o$	0.070	-0.009^{***}	$+0.010^{***}$	-0.006^{***}
$r \to o$	0.092	-0.002^{**}	$+0.004^{***}$	-0.008^{***}
$m \to o$	0.143	$+0.023^{***}$	$+0.029^{***}$	-0.013^{***}
(Re-)entries from non-employment				
$o \to m$	0.057	-0.042^{***}	$+0.015^{***}$	-0.006^{***}
$o \to r$	0.090	-0.047^{***}	-0.011^{***}	-0.016^{***}
$o \to a$	0.029	$+0.071^{***}$	-0.019^{***}	-0.004^{***}

Notes: Each row reports coefficients from a separate LPM regression of the indicated transition on a college dummy, a female dummy, age, tenure, and year fixed effects, with each year weighted equally. $\bar{y}$ is the unconditional probability of the transition within the origin-state sample. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; full regression tables are in Appendix B.4.

through entrants and the less-tenured rather than through reshuffling of long-tenured incumbents. Age coefficients are small once tenure is included and are omitted from the table; the visible exceptions are $a \to o$ and $r \to o$, where older workers are more likely to exit to non-employment (end-of-career retirement) even as shorter tenure independently raises exit risk.

Third, *women upgrade less and exit or downgrade more.* They are penalized on all three upgrading margins ($r \to a$, -2.2 pp; $m \to r$, -3.8; $m \to a$, -0.9) and are less likely to re-enter the two higher destinations from non-employment ($o \to r$, -1.1; $o \to a$, -1.9). At the same time they are more likely to exit to non-employment from every origin ($m \to o$, $+2.9$; $r \to o$, $+0.4$; $a \to o$, $+1.0$) and more likely to downgrade from abstract to routine ($a \to r$, $+3.7$). The only entry margin where they exceed men is re-entry into manual ($o \to m$, $+1.5$).

5.2.2 Across the 1994 Break

We now re-estimate the same specification separately on the pre-polarization (1984–1993) and polarization (1994–2021) sub-samples (Table 5) to explore the extent to which the selection into transitions is stable across the break. We observe that the previous three patterns are indeed regime-dependent.

First, women's penalty is reduced across 1994 on every margin. The female disadvantage on $r \to a$ upgrading falls from -2.6 to -2.0 pp and on $m \to r$ from -6.5 to -2.9 pp, and women's excess probability of downgrading $a \to r$ shrinks from $+5.1$ to $+3.1$ pp. The largest move is on routine re-entry: the female coefficient on $o \to r$ contracts from -2.8 to -0.3 pp, so the gender gap there nearly closes. With the routine re-entry channel contracting overall (Section 4), this suggests that the routine entry contraction is male-driven.

Second, the college premium on upgrading roughly halves across the break — on $r \to a$ from $+20.2$ to $+8.1$ pp, on $m \to a$ from $+11.8$ to $+7.0$ pp — and the college coefficient on $o \to r$ halves as well ($-7.8 \to -4.0$ pp). Upgrading becomes less selective on education during polarization, consistent with the Section 5.1 findings that the college routine-to-abstract net flow more than halves while the non-college pool still accounts for most of those flows.

Third, tenure's stabilizing role appears to weaken across the break: the employed-origin tenure coefficients fall to roughly a third of their pre-1994 size — $m \to r$ from -3.6 to -1.0 pp per year of tenure, $a \to r$ from -2.8 to -0.9 — and the non-employment-duration coefficient on routine re-entry falls even more sharply ($o \to r$, -8.0 to -1.3).[18]

[18] This cross-period comparison is partly a panel-length artefact: tenure is bounded by how long the panel has run, so before 1994 within-CS2 tenure cannot exceed 9 years (1984–1993), against 26 in the polarization sub-sample. Appendix B.4.1 re-estimates the split with tenure capped at nine years in both periods (Table 35): for the employed-origin transitions the polarization-to-pre-polarization coefficient ratio rises from about 0.4 to 0.7, so most of the apparent weakening is mechanical. The exception is the

Table 5: Partial correlations by transition, split at the 1994 break: pre-polarization (1984–1993) and polarization (1994–2021)

	Pre-polarization (1984–93)			**Polarization (1994–2021)**		
Transition	College	Female	Tenure	College	Female	Tenure
Within-employment upgrading						
$m \to r$	$+0.030^{*}$	-0.065^{***}	-0.0358^{***}	$+0.061^{***}$	-0.029^{***}	-0.0096^{***}
$r \to a$	$+0.202^{***}$	-0.026^{***}	-0.0065^{***}	$+0.081^{***}$	-0.020^{***}	-0.0029^{***}
$m \to a$	$+0.118^{***}$	-0.009^{***}	-0.0024^{***}	$+0.070^{***}$	-0.009^{***}	-0.0012^{***}
Within-employment downgrading						
$r \to m$	-0.020^{***}	-0.003^{***}	-0.0086^{***}	-0.014^{***}	-0.004^{***}	-0.0034^{***}
$a \to r$	-0.098^{***}	$+0.051^{***}$	-0.0278^{***}	-0.055^{***}	$+0.031^{***}$	-0.0086^{***}
$a \to m$	-0.007^{***}	$+0.006^{***}$	-0.0026^{***}	-0.008^{***}	$+0.000$	-0.0010^{***}
Exits to non-employment						
$m \to o$	$+0.056^{***}$	$+0.045^{***}$	-0.0189^{***}	$+0.018^{***}$	$+0.023^{***}$	-0.0128^{***}
$r \to o$	$+0.024^{***}$	$+0.016^{***}$	-0.0154^{***}	-0.005^{***}	$+0.000$	-0.0079^{***}
$a \to o$	-0.007^{*}	$+0.020^{***}$	-0.0176^{***}	-0.009^{***}	$+0.006^{***}$	-0.0057^{***}
(Re-)entries from non-employment						
$o \to m$	-0.035^{***}	$+0.002^{*}$	-0.0213^{***}	-0.045^{***}	$+0.020^{***}$	-0.0050^{***}
$o \to r$	-0.078^{***}	-0.028^{***}	-0.0804^{***}	-0.040^{***}	-0.003^{***}	-0.0134^{***}
$o \to a$	$+0.069^{***}$	-0.026^{***}	-0.0181^{***}	$+0.072^{***}$	-0.017^{***}	-0.0034^{***}

Notes: Coefficients (in probability units) from the specification of Table 4, estimated separately on each sub-period with each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

The emerging picture is that polarization changed the selection into the flows, not just their size. The characteristics that sort workers into upgrading and into routine re-entry are regime-dependent, shifting at the break — which the pooled estimates average away.

6 Conclusion

Using French administrative panel data, we show that labor-market polarization emerges mainly from changing entry dynamics rather than mass displacement of incumbent routine workers. We find that the decisive change after 1994 is a collapse in the net flow from non-employment into routine: non-employment ceases to be a net feeder of the middle-skill routine occupations and the entry collapse accounts for about four-fifths of the post-1994 acceleration of the routine decline. We also find that the rise in manual employment is driven mainly by entrants from non-employment disproportionately flowing into manual jobs and, to a lesser extent, by a dampening of the net upgrading flows to routine. All along, routine to abstract upgrading flows are large and stable throughout 1984–2021.

Notably, this upgrading pattern is not restricted to college-educated workers: non-college workers account for almost half of the rise in the abstract share through their own substantial routine-to-abstract upgrading. The welfare burden of structural change therefore concentrates on new entrants lacking education, rather than on incumbent routine workers, who tend to upgrade to abstract at constant rates.

All in all, this behavior of worker flows clearly points to the entry margin. Thus, if labor market policy is meant to ameliorate the impact of polarization, then expanding or retooling education and training and easing geographic mobility are likely to be more effective than policies aimed primarily at protecting incumbent workers from displacement. Nonetheless, the impact on welfare and the consequences of policies need to be evaluated using a quantitative life-cycle model of occupational choice and entry and exit from labor markets. In a companion paper we calibrate such a model to these transition flows to assess cohort welfare and the relative roles of technology and demographic changes. (Cerina et al., 2026a)

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4B, pages 1043–1171. Elsevier.

Acemoglu, D. and Loebbing, J. (2026). Automation and polarization. *Journal of Political Economy*. Forthcoming.

re-entry margin, where the attenuation survives capping (ratios below 0.3) and is genuine.

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

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

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

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

Bocquet, L. (2024). The network origin of slow labor reallocation. Paris School of Economics / Cambridge Working Paper.

Bozio, A. and Wasmer, É. (2024). Les politiques d'exonérations de cotisations sociales : une inflexion nécessaire. Rapport au premier ministre, Haut-commissariat à la stratégie et au plan, Paris.

Cerina, F., Dienesch, E., Monge-Naranjo, A., and Moro, A. (2026a). On the mechanics of labor market polarization. Mimeo, Università di Cagliari and Federal Reserve Bank of Atlanta.

Cerina, F., Dienesch, E., Monge-Naranjo, A., and Moro, A. (2026b). Urban elevators and the geography of labor market polarization. Mimeo, Università di Cagliari and Federal Reserve Bank of Atlanta.

Cerina, F., Dienesch, E., Moro, A., and Rendall, M. (2023). Spatial polarisation. *Economic Journal*, 133(649):30–69.

Cerina, F., Moro, A., and Rendall, M. (2021). The role of gender in employment polarization. *International Economic Review*, 62(4):1655–1691.

Charlot, O., Fontaine, I., and Sopraseuth, T. (2024). Job polarization and non-standard work: Evidence from France. *Labour Economics*, 88:102534.

Cortes, G. M. (2016). Where have the middle-wage workers gone? A study of polarization using panel data. *Journal of Labor Economics*, 34(1):63–105.

Cortes, G. M., Jaimovich, N., Nekarda, C. J., and Siu, H. E. (2020). The dynamics of disappearing routine jobs: A flows approach. *Labour Economics*, 65:101823.

Cortes, G. M., Jaimovich, N., and Siu, H. E. (2017). Disappearing routine jobs: Who, how, and why? *Journal of Monetary Economics*, 91:69–87.

Davis, D. R., Mengus, E., and Michalski, T. K. (2026). Labor market polarization and the great urban divergence. *Journal of International Economics*, 160:104224.

García-Peñalosa, C., Petit, F., and van Ypersele, T. (2023). Can workers still climb the social ladder as middling jobs become scarce? Evidence from two British cohorts. *Labour Economics*, 84:102390.

Garibaldi, P., Gomes, P., and Sopraseuth, T. (2025). Economic consequences of vertical mismatch. *Quantitative Economics*, 16(2):535–564.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, 104(8):2509–2526.

Jaimovich, N. and Siu, H. E. (2020). Job polarization and jobless recoveries. *Review of Economics and Statistics*, 102(1):129–147.

Kashkarov, D. and Artemev, V. (2024). Disappearing stepping stones: Technological change and career paths. Working Paper.

A Sample Description Details

A.1 Wage Structure by CS2 Code

Table 6 ranks all CS2 codes by mean daily net wage in four benchmark years, expressed relative to the manual-group average ($\bar{w}_m = 1$). The sample is restricted to employed workers aged 25–64 with positive wages.

Abstract pays roughly 3× manual throughout the period, routine cognitive roughly 1.5–1.7×, and routine manual roughly 1.2–1.5×. We note that *rc* and *rm* occupy distinct ranges of the distribution: *rc* codes span ranks 5–21 (overlapping with the bottom of abstract), while *rm* codes cluster in ranks 10–22 (overlapping with the top of manual).

A.2 Employment Spells and NE Gaps

An *employment spell* is a maximal sequence of consecutive years in which the worker is employed (in one of m, r, a). An *NE interruption gap* is a maximal sequence of consecutive years of non-employment (o), flanked by employment on both sides.

Among the 1.86 million ever-employed workers in the panel, 51.5% have a single continuous employment spell and 48.5% have multiple spells (interrupted by NE gaps). The median employment spell lasts 3 years; the distribution is right-skewed, with 26% of spells lasting only one year and a long right tail extending past 14 years.

NE interruption gaps are mostly short: 43% last one year, 23% last two years, and the median gap is 2 years. Longer gaps ($\geq$ 5 years) account for 14% of all gaps and likely include exits to self-employment, long-term unemployment, or return migration not captured in the panel.

A.3 Age-Boundary Analysis

The flow matrix F uses the 25–64 age restriction at $t-1$ only: a worker is in the flow sample for the ($t-1$, t) pair if and only if she is aged 25–64 at $t-1$, regardless of her age at t. The employment stock $N_k(t)$, by contrast, applies the 25–64 age restriction at t. This asymmetry generates two boundary cohorts. *Aging-in* workers are aged 24 at $t-1$ and 25 at t: they are excluded from the flow sample (they fail the $t-1$ age cut) but appear in the stock $N_k(t)$. *Aging-out* workers are aged 64 at $t-1$ and 65 at t: they appear in the flow sample but are excluded from $N_k(t)$.

To preserve the accounting identity in equation (5), we code aging-in workers (employed and aged 25–64 at t, but outside the age window at $t-1$) as $o \to k$ entries during period t, and aging-out workers (employed at $t-1$, outside the age window at t) as $k \to o$ exits. With this convention the recovered $F_{o,k}(t)$ in equation (6) absorbs aging-in flows automatically: they enter $\Delta N_k(t)$ through the t snapshot but contribute to neither JtJ

Table 6: CS2 codes ranked by mean daily wage relative to manual average

CS2	Description	Group	1984 Rk	1984 $w/\bar{w}_m$	1994 Rk	1994 $w/\bar{w}_m$	2009 Rk	2009 $w/\bar{w}_m$	2019 Rk	2019 $w/\bar{w}_m$
35	Prof. information/arts	a	1	6.29	7	2.37	9	2.05	4	2.42
37	Cadres admin./comm.	a	2	3.68	2	4.19	2	3.50	2	3.35
38	Ingénieurs/cadres tech.	a	3	3.52	4	3.62	3	3.33	3	3.00
23	Chefs entr. $\geq$10 sal.	a	4	3.28	1	5.01	1	5.43	1	3.81
33	Cadres fonc. publique	a	–	–	6	2.51	5	2.41	6	2.17
47	Techniciens	a	7	1.89	8	2.32	10	2.05	8	1.85
34	Prof. enseign./santé	a	8	1.76	3	3.71	6	2.39	5	2.37
31	Professions libérales	a	–	–	–	–	4	2.45	9	1.84
	abstract average	a		3.00		3.39		2.88		2.74
46	Prof. int. admin./comm.	rc	5	2.01	10	1.90	11	1.94	10	1.80
48	Contremaîtres	rc	6	1.98	9	2.09	7	2.19	7	1.99
43	Prof. int. santé/social	rc	9	1.58	5	2.74	12	1.78	13	1.56
54	Empl. admin. entreprise	rc	12	1.33	15	1.39	18	1.46	18	1.37
45	Prof. int. fonc. publ.	rc	–	–	12	1.64	8	2.10	11	1.68
42	Instituteurs	rc	14	1.26	21	0.90	14	1.65	15	1.48
52	Empl. civils fonc. publ.	rc	15	1.19	14	1.44	20	1.40	21	1.22
	rc average	rc		1.59		1.71		1.64		1.50
65	Ouv. qual. manut./transp.	rm	10	1.43	11	1.80	16	1.58	14	1.48
62	Ouv. qual. industriel	rm	13	1.32	13	1.60	13	1.69	16	1.47
63	Ouv. qual. artisanal	rm	16	1.17	16	1.39	17	1.49	19	1.36
67	Ouv. non qual. industriel	rm	18	1.05	19	1.29	22	1.27	22	1.17
	rm average	rm		1.22		1.47		1.52		1.39
	routine average (r)	r		1.41		1.62		1.60		1.46
64	Chauffeurs	m	11	1.39	18	1.34	19	1.41	20	1.34
53	Policiers/militaires	m	17	1.13	17	1.36	15	1.59	12	1.58
55	Empl. de commerce	m	19	1.00	20	1.05	24	1.15	24	1.13
56	Pers. services partic.	m	20	0.84	24	0.78	26	0.66	26	0.71
68	Ouv. non qual. artisanal	m	21	0.78	23	0.80	25	0.89	25	0.83
69	Ouv. non qual. divers	m	22	0.68	22	0.89	23	1.21	23	1.15
10	Agriculteurs	m	–	–	–	–	21	1.39	17	1.43
	manual average	m		1.00		1.00		1.00		1.00

	Group rank		1984		1994		2009		2019	
	abstract	a	1	3.00	1	3.39	1	2.88	1	2.74
	routine cognitive	rc	2	1.59	2	1.71	2	1.64	2	1.50
	routine (all)	r	2	1.41	2	1.62	2	1.60	2	1.46
	routine manual	rm	3	1.22	3	1.47	3	1.52	3	1.39
	manual	m	4	1.00	4	1.00	4	1.00	4	1.00

Note: Mean daily net wage by CS2 code, workers with positive wages. Rk = rank among all CS2 codes within that year. $w/\bar{w}_m$ = ratio of CS2 mean wage to the manual-group average. Group averages are worker-weighted. Dashes indicate fewer than 10 observations (small pre-2002 sample).

nor $F_{k,o}$, so they pass through into the residual. Aging-out flows enter the observed exit count $F_{k,o}(t)$ directly. The identity therefore holds without further adjustment.

This convention underlies the *direct tabulation* method of Appendix C.2: with boundary-crossers coded as $o \to k$ entries or $k \to o$ exits, it reproduces the baseline's JtJ flows, exits, and entries for every one-year pair, differing only in the $o \to o$ cell, which does not enter the employment-share decomposition.

Workers first entering the 25–64 window at age 25 number 468,000 (25.2% of all ever-employed workers). Their initial occupation composition is manual-overweight (26.2% vs. 20.5% in the full population) and abstract-underweight (15.4% vs. 22.1%). This composition is roughly constant over time, so the demographic bias shifts the *level* of non-employment-mediated flows but does not generate a spurious trend. Workers aging out at 64 are an order of magnitude smaller (69,000 workers); their composition is manual-overweight (33.4% vs. 20.5%) and abstract-overweight (29.0% vs. 22.1%), with routine correspondingly underweight (37.6% vs. 57.4%).

Boundary flows by construction affect only the non-employment channel: JtJ flows require both origin and destination to be in employed states (in $\{m, r, a\}$), so age-boundary composition cannot enter them. This is confirmed empirically: JtJ bilateral flows are virtually identical in the 30–59 and 25–64 windows (differences below 0.01 pp/year for every bilateral pair in every sub-period; see Appendix C.1).

B Additional Evidence

B.1 Full Transition Matrices

This appendix reports the simple-average conditional and unconditional transition matrices for the three periods discussed in the main text. Each row of a conditional matrix gives the probability of being in state κ' at $t+1$ conditional on being in state κ at t; rows sum to one. Each cell of an unconditional matrix gives the probability of the (κ, κ') transition in the population; the full matrix sums to one.

B.1.1 Conditional Matrices

Table 7: Conditional matrix — full period 1984–2021.

From κ \ To κ'	o	m	r	a
o	0.690	0.085	0.182	0.043
m	0.161	0.721	0.108	0.010
r	0.107	0.030	0.830	0.033
a	0.093	0.008	0.085	0.814

Table 8: Conditional matrix — pre-polarization 1984–1993.

From κ \ To κ'	o	m	r	a
o	0.662	0.072	0.225	0.040
m	0.193	0.634	0.163	0.010
r	0.138	0.037	0.786	0.038
a	0.134	0.010	0.140	0.717

Table 9: Conditional matrix — polarization 1994–2021.

From κ \ To κ'	o	m	r	a
o	0.704	0.088	0.164	0.044
m	0.149	0.753	0.089	0.009
r	0.096	0.027	0.846	0.031
a	0.078	0.007	0.067	0.848

The key regime change is visible in the $o \to r$ cell: from 0.225 in the pre-polarization phase to 0.164 during polarization, a 27% decline. $o \to m$ instead rises ($0.072 \to 0.088$), so the entry-margin collapse is specific to routine. All diagonal entries increase from pre-polarization to polarization (persistence rises), consistent with lower overall mobility. The $r \to a$ upgrading probability is remarkably similar across regimes (0.038 pre-pol, 0.031 pol).

B.1.2 Unconditional (Joint) Matrices

Table 10: Unconditional matrix — full period 1984–2021

From κ \ To κ'	o	m	r	a	Row sum
o	0.206	0.025	0.055	0.013	0.299
m	0.021	0.097	0.014	0.001	0.134
r	0.047	0.013	0.358	0.014	0.431
a	0.012	0.001	0.011	0.112	0.135
Col sum	0.286	0.136	0.437	0.140	1.000

The unconditional matrices reveal the population weight of each flow—the joint flows that the employment-share decomposition of Section 4 aggregates. The $o \to r$ cell drops from 0.074 (pre-pol) to 0.048 (pol), the largest off-diagonal decline in the matrix. The r row sum (which reflects routine's population share) declines from 0.456 to 0.422 across the two regimes, but the composition of flows out of routine shifts markedly: exits to

Table 11: Unconditional matrix — pre-polarization 1984–1993

From κ \ To κ'	o	m	r	a	Row sum
o	0.219	0.024	0.074	0.013	0.331
m	0.022	0.071	0.018	0.001	0.112
r	0.063	0.017	0.359	0.018	0.456
a	0.014	0.001	0.014	0.073	0.101
Col sum	0.317	0.113	0.465	0.105	1.000

Table 12: Unconditional matrix — polarization 1994–2021

From κ \ To κ'	o	m	r	a	Row sum
o	0.203	0.025	0.048	0.013	0.288
m	0.021	0.107	0.012	0.001	0.141
r	0.040	0.012	0.357	0.013	0.422
a	0.011	0.001	0.009	0.126	0.148
Col sum	0.276	0.145	0.427	0.153	1.000

non-employment fall (from 0.063 to 0.040) while the diagonal rises, consistent with higher within-state persistence among the shrinking routine population.

B.2 Additional Evidence on Occupational Flows

This appendix collects the fine sub-period decomposition that refines the pre-polarization and polarization regimes, the five-state decomposition with routine split into cognitive and manual sub-groups, the analysis of two-step transitions via non-employment, and an entry-vs-exit anatomy of the routine non-employment channel.

B.2.1 Fine Sub-period Decomposition

Table 13 splits the polarization regime reported in Table 2 of the main text into three non-overlapping windows: a pre-crisis phase (1994–2006), a post-crisis phase (2009–2019), and the polarization period excluding the crisis and pandemic boundary years (1994–2019 excluding 2007–2009).

Three facts emerge. First, the $r \to a$ bilateral flow is roughly constant at about -0.50 pp/year in every sub-period, so the upgrading drain is a structural feature of the French labor market rather than a polarization-era phenomenon. Second, the pre-crisis phase 1994–2006 displays both the weakening of the $m \to r$ buffer (from +0.20 pre-polarization to +0.08) and the collapse of the routine Excess_NE channel (from +0.25 to

Table 13: Detailed bilateral decomposition by sub-period (pp/year)

Period		Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
Pre-crisis 1994–2006	manual	—	−0.08	−0.03	−0.11	+0.34	+0.23
	routine	+0.08	—	−0.49	−0.41	−0.03	−0.44
	abstract	+0.03	+0.49	—	+0.52	−0.31	+0.21
Post-crisis 2009–2019	manual	—	−0.14	−0.07	−0.21	+0.12	−0.09
	routine	+0.14	—	−0.49	−0.35	+0.08	−0.27
	abstract	+0.07	+0.49	—	+0.56	−0.20	+0.36
Polarization excl. 2007–09 1994–2019	manual	—	−0.11	−0.05	−0.16	+0.23	+0.07
	routine	+0.11	—	−0.49	−0.38	+0.03	−0.36
	abstract	+0.05	+0.49	—	+0.54	−0.26	+0.28

Notes: Values ×100 (pp/year). Same format as Table 2 in the main text. The bottom panel excludes pairs with $t_0 \in \{2007, 2008\}$ (crisis).

−0.03) in their cleanest form; among these three sub-periods it has the steepest routine decline ($\Delta E_r = -0.44$). Third, the partial post-crisis recovery of both buffers explains why the full 1994–2021 Excess_NE for routine in Table 2 is more negative than the sub-period averages taken separately: it absorbs the crisis outliers of 2007–09. Excluding those pairs, the routine Excess_NE channel is close to zero rather than a clean sign reversal, as noted in the main text.

B.2.2 Five-state Decomposition with RC/RM Split

Table 14 reports the five-state mean annual decomposition underlying Figure 6 in the main text. Pre-polarization and polarization sub-periods match those used in Table 2; states rc and rm denote the routine cognitive and routine manual sub-groups defined in Section 4.4.

Table 14: Five-state mean annual decomposition (pp/year).

	Pre-pol 1984–1993			**Pol 1994–2021**			**Full 1984–2021**		
State	Net_JtJ	Excess_NE	ΔE	Net_JtJ	Excess_NE	ΔE	Net_JtJ	Excess_NE	ΔE
manual	−0.24	+0.04	−0.20	−0.13	+0.25	+0.12	−0.16	+0.19	+0.03
rc	−0.03	+0.80	+0.77	−0.28	+0.04	−0.25	−0.22	+0.24	+0.02
rm	−0.31	−0.56	−0.86	−0.15	−0.12	−0.27	−0.19	−0.24	−0.43
abstract	+0.58	−0.29	+0.29	+0.57	−0.16	+0.40	+0.57	−0.19	+0.37

Notes: Values ×100 (pp/year). Five-state system splits routine into routine cognitive (rc, CS2 $\in \{42, 43, 45, 46, 48, 52, 54\}$) and routine manual ($rm$, CS2 $\in \{62, 63, 65, 67\}$).

B.2.3 Non-Employment as an Upgrading Device

A natural concern about the Excess_NE channel is that it may conceal *disguised occupational downgrading*: routine workers who lose their job, spend one period in non-employment, and then re-enter as manual workers would appear in the accounts as an $r \to o$ exit followed by an $o \to m$ entry, not as a direct $r \to m$ downgrade. If the indirect $r \to o \to m$ path were quantitatively important, the upgrading story documented in Section 4 could be an artefact of focusing on direct JtJ flows only. To address this concern we construct two-step transitions via non-employment: for each year-triplet $(t, t{+}1, t{+}2)$ we identify workers employed at t, non-employed at $t{+}1$, and re-employed at $t{+}2$, and compute the 3×3 origin×destination matrix of their employed-to-employed outcomes. Simple-averaging across the 33 triplets gives the unconditional via-o matrix in Table 15.

Table 15: Unconditional via-o matrix: origin(t)× destination($t{+}2$) for workers employed at t, non-employed at $t{+}1$, and re-employed at $t{+}2$. Full period

	$\to m$	$\to r$	$\to a$	Row sum
manual	0.201	0.109	0.015	0.325
routine	0.099	0.418	0.046	0.563
abstract	0.007	0.029	0.077	0.112
Column sum	0.307	0.555	0.138	1.000

Notes: Aggregated across 33 year-triplets $(t, t{+}1, t{+}2)$ covering 1984–2019. Each cell is the share of via-NE workers following origin $\to o \to$ destination. Diagonal cells are bouncebacks (origin = destination).

Three features are worth highlighting. First, the diagonal dominates: workers who pass through non-employment overwhelmingly return to the same broad occupation ($r \to o \to r$ equals 0.42, $m \to o \to m$ equals 0.20, $a \to o \to a$ equals 0.08). To a first approximation, non-employment is an absorbing state within occupation rather than a channel of reallocation. Second, every bilateral comparison is *net upgrading* rather than net downgrading: $m \to o \to r$ (0.11) exceeds $r \to o \to m$ (0.10); $m \to o \to a$ (0.015) exceeds its reverse (0.007); and $r \to o \to a$ (0.046) exceeds $a \to o \to r$ (0.029). Even through the indirect o channel, net flows are upgrading. The disguised-downgrading concern is therefore not supported by the data. Third, the pattern intensifies over time. Figure 11 plots the three net bilateral via-o flows by year: after the 2008 crisis, all three bilateral pairs settle into a regime of stable net upgrading, and the r-to-a via-o gap widens as $a \to o \to r$ falls while $r \to o \to a$ remains stable.

Non-employment is therefore not primarily a waystation for displaced routine workers descending to manual, but rather a search interval through which workers on average emerge at a higher rung of the occupational ladder than at a lower one. When we widen the window to include indirect paths through non-employment, the upgrading asymmetry

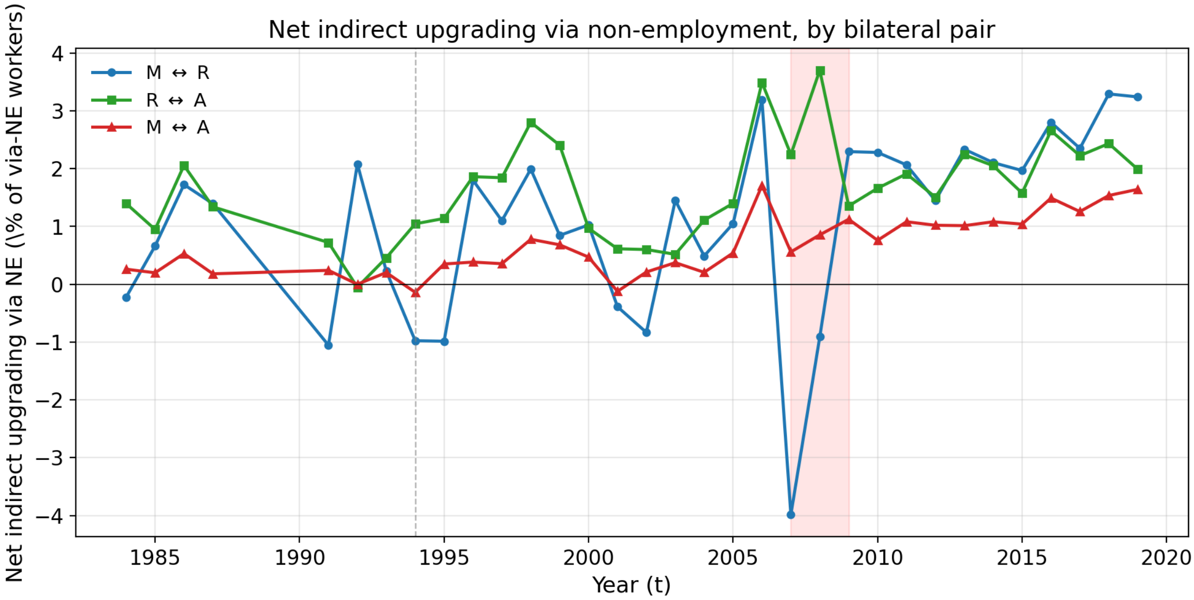

Figure 11: Net indirect upgrading via non-employment, by bilateral pair and year. For each pair (k, ℓ), the line reports the unconditional share of via-o workers following $k \to o \to \ell$ minus the share following $\ell \to o \to k$. Positive values indicate net upgrading.

documented in Section 4 for direct JtJ flows is preserved, not reversed.

B.2.4 Additional evidence on the education margin

This appendix collects supplementary evidence underlying the education-stratified results of Section 5.1. Subsection B.2.4 formalises the Oaxaca-Blinder decomposition of within-employed share changes by education group. Subsection B.2.4 reports the bilateral Net_JtJ decompositions by education group for three windows: the full panel 1984–2021, the pre-polarization period 1984–1993, and the polarization period 1994–2021.

Oaxaca-Blinder decomposition of share changes For each occupational state $k \in \{m, r, a\}$, the aggregate within-employed share decomposes by education as

$$E_k(t) \;=\; w_c(t) \cdot E_k^c(t) \;+\; w_{nc}(t) \cdot E_k^{nc}(t),$$

where $w_c(t)$ is the college share among employed workers with education observed, $w_{nc}(t) = 1 - w_c(t)$, and $E_k^c(t)$ and $E_k^{nc}(t)$ are the within-employed shares of state k inside the college and non-college groups respectively.

Differencing between two endpoints and rearranging yields the three-way decomposition

$$\Delta E_k \;=\; \underbrace{\bar{w}_{nc}\,\Delta E_k^{nc}}_{\text{within non-college}} \;+\; \underbrace{\bar{w}_c\,\Delta E_k^c}_{\text{within college}} \;+\; \underbrace{\Delta w_c \cdot (\bar{E}_k^c - \bar{E}_k^{nc})}_{\text{composition}},$$

where bars denote midpoint averages of the two endpoints. The first two terms capture changes within each education group at fixed group weights; the third term captures the contribution of the shifting college share of employment at fixed within-group shares.

Table 16 reports the decomposition for each occupational state over the full panel and the pre-polarization and polarization sub-periods.

Table 16: Oaxaca-Blinder decomposition of ΔE_k by education group. Values in percentage points (midpoint weights).

		Total	Within non-coll.	Within coll.	Composition
Manual	Full 1984–2021	+4.8	+6.8	−0.1	−1.9
	Pre-pol 1984–1993	+0.3	+0.5	−0.1	−0.2
	Pol 1994–2021	+5.4	+7.0	+0.1	−1.6
Routine	Full 1984–2021	−16.3	−11.8	−0.8	−3.8
	Pre-pol 1984–1993	−2.8	−1.7	−0.5	−0.6
	Pol 1994–2021	−13.5	−10.3	+0.1	−3.3
Abstract	Full 1984–2021	+11.5	+5.0	+0.8	+5.7
	Pre-pol 1984–1993	+2.5	+1.2	+0.5	+0.8
	Pol 1994–2021	+8.1	+3.3	−0.2	+4.9

Notes: Each row decomposes the aggregate within-employed share change ΔE_k into within-non-college, within-college, and composition (rising college share) components, using the Oaxaca-Blinder identity with midpoint weights (see Appendix B.2.4). Sample restricted to employed workers with education observed; college share among employed-with-education rises from 6.0% in 1984 to 17.7% in 2021.

Three observations summarise the table. First, in every state and period the within-non-college component is the dominant within-group contributor; for the abstract gain, the within-non-college contribution alone is roughly half of the aggregate rise. Second, the rising college share (composition term) is a non-trivial contributor to the abstract and routine changes: because the college sub-population has substantially more abstract employment and less routine employment than the non-college sub-population, the rising college share mechanically lifts the aggregate abstract share and depresses the aggregate routine share. Third, the aggregate manual rise (+4.8 pp) reflects a substantially larger within-non-college increase (+6.8 pp), partly offset by the compositional shift toward the (lower-manual) college pool (−1.9 pp); the within-college trajectory is essentially flat.

Robustness to weight scheme (start, midpoint, end weights), to endpoint choice, and to restricting the sample to post-2002 worker-years where education is genuinely observed (rather than back-projected) gives a robust non-college share of the abstract rise in the range 39–47% (midpoint 43%).

Bilateral flows by education and period Tables 17–19 report the simple-average bilateral Net_JtJ decomposition by education group for the three windows. Each table contains three sub-blocks: All workers with education observed, Non-college, and College.

For each origin state $k \in \{m, r, a\}$ the table reports the three bilateral net flows into k from each origin (Net-k' columns), their sum (Net_JtJ), the residual Excess_NE channel, and the resulting ΔE_k. Reading down a sub-block, the three rows sum to zero up to rounding.

Table 17: Mean annual bilateral decomposition by education group, full panel 1984–2021 (pp/year).

Group	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	manual	$---$	-0.13	-0.04	-0.18	$+0.24$	$+0.06$
	routine	$+0.13$	$---$	-0.53	-0.39	-0.00	-0.40
	abstract	$+0.04$	$+0.53$	$---$	$+0.57$	-0.24	$+0.33$
Non-college	manual	$---$	-0.13	-0.04	-0.17	$+0.31$	$+0.14$
	routine	$+0.13$	$---$	-0.45	-0.32	$+0.00$	-0.32
	abstract	$+0.04$	$+0.45$	$---$	$+0.49$	-0.31	$+0.18$
College	manual	$---$	-0.13	-0.10	-0.23	$+0.19$	-0.04
	routine	$+0.13$	$---$	-1.27	-1.14	$+1.02$	-0.12
	abstract	$+0.10$	$+1.27$	$---$	$+1.37$	-1.21	$+0.16$

Notes: Values $\times 100$ (pp/year), sample restricted to workers with education observed (worker-level forward/backward fill). Net-k' entries report the bilateral net JtJ flow into k from origin k'.

Table 18: Mean annual bilateral decomposition by education group, pre-polarization period 1984–1993 (pp/year).

Group	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	manual	$---$	-0.31	-0.04	-0.35	$+0.20$	-0.15
	routine	$+0.31$	$---$	-0.58	-0.27	$+0.17$	-0.10
	abstract	$+0.04$	$+0.58$	$---$	$+0.62$	-0.37	$+0.26$
Non-college	manual	$---$	-0.32	-0.04	-0.36	$+0.23$	-0.12
	routine	$+0.32$	$---$	-0.47	-0.14	$+0.17$	$+0.02$
	abstract	$+0.04$	$+0.47$	$---$	$+0.50$	-0.40	$+0.10$
College	manual	$---$	-0.10	-0.12	-0.22	$+0.13$	-0.09
	routine	$+0.10$	$---$	-2.17	-2.07	$+1.62$	-0.45
	abstract	$+0.12$	$+2.17$	$---$	$+2.29$	-1.75	$+0.53$

Notes: Values $\times 100$ (pp/year). Sample as in Table 17.

Two patterns reinforce the comparison reported in Section 5.1. First, the bilateral $r \to a$ flow inside the non-college sample is essentially unchanged across the two regimes (+0.47 pp/year pre-polarization vs +0.45 post-1994), while inside the college sample it is large in both periods (+2.17 vs +0.98). The non-college routine decline during polarization therefore reflects the weakening of the $m \to r$ buffer flow (+0.32 to +0.07) and the sign flip of the non-employment entry channel into routine (+0.17 to −0.05), not a slowdown in

Table 19: Mean annual bilateral decomposition by education group, polarization period 1994–2021 (pp/year).

Group	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	manual	$---$	-0.08	-0.04	-0.12	$+0.25$	$+0.13$
	routine	$+0.08$	$---$	-0.51	-0.43	-0.06	-0.49
	abstract	$+0.04$	$+0.51$	$---$	$+0.55$	-0.20	$+0.36$
Non-college	manual	$---$	-0.07	-0.03	-0.11	$+0.33$	$+0.22$
	routine	$+0.07$	$---$	-0.45	-0.38	-0.05	-0.43
	abstract	$+0.03$	$+0.45$	$---$	$+0.48$	-0.28	$+0.20$
College	manual	$---$	-0.14	-0.10	-0.24	$+0.21$	-0.03
	routine	$+0.14$	$---$	-0.98	-0.84	$+0.82$	-0.02
	abstract	$+0.10$	$+0.98$	$---$	$+1.08$	-1.03	$+0.05$

Notes: Values $\times 100$ (pp/year). Sample as in Table 17.

$r \to a$ upgrading. Second, the manual share rise during polarization inside the non-college sample ($\Delta E_m = +0.22$ pp/year, vs -0.13 pre-polarization) runs against an interpretation as displacement from routine into manual: the bilateral $r \to m$ flow runs in the opposite direction in both periods (the $m \to r$ flow is positive at $+0.32$ and $+0.07$ respectively). The non-college manual rise is instead absorbed by the manual non-employment entry channel (Excess_NE$_m$ = $+0.33$ pp/year during polarization).

B.3 Supplementary Material on the RC / RM Split

Section 4.4.1 introduces the routine-cognitive (rc) / routine-manual (rm) disaggregation of the routine category and reports the sub-period flow decomposition used to demonstrate that the aggregate routine Excess_NE sign flip is an rc phenomenon. This appendix extends the related evidence by presenting the o →CS2 entry-rate breakdown by destination and the full five-state unconditional transition matrices.

B.3.1 Entry Rates by Destination

Table 20 reports the o →CS2 entry rate (annual average o →CS2 transitions divided by the non-employed population) by destination CS2, aggregated within the rc and rm groups. The entry-rate decline from pre-polarization to polarization is broad-based: rc loses -2.37 pp of o-entry inflow (a 32% drop) and rm loses -1.51 pp (34%). In percentage terms the two groups decline comparably; in absolute pp terms the decline is larger for rc because rc was the larger destination pre-polarization. Importantly, both white-collar rc codes (CS46-54) and blue-collar rm codes (CS62-67) experience sizeable entry-rate drops, ruling out a pure deindustrialization reading of the entry-margin collapse.

Table 20: $o \to$CS2 entry rate by destination (% of o population), aggregated into rc and rm groups

CS2	Description	Pre-pol 84–93	Pol 94–06	Post-crisis 09–19
Routine cognitive (rc)				
42	Instituteurs	0.42	0.30	0.43
43	Prof. interm. santé	0.68	0.50	0.50
45	Prof. interm. admin. publ.	0.29	0.13	0.08
46	Prof. interm. admin. entr.	1.68	1.12	0.69
48	Contremaîtres	0.48	0.25	0.12
52	Empl. civils fonct. publ.	1.74	1.32	1.37
54	Empl. admin. d'entreprise	2.19	1.50	1.31
	rc total	**7.48**	**5.11**	**4.50**
Routine manual (rm)				
62	Ouvr. qual. type artisanal	1.47	0.76	0.85
63	Ouvr. qual. type industriel	1.25	1.05	1.04
65	Ouvr. qual. manut., magas.	0.30	0.21	0.26
67	Ouvr. non qual. type indus.	1.67	1.09	0.89
	rm total	**4.69**	**3.11**	**3.05**

Notes: Annual $o \to$CS2 entry transitions divided by the non-employed population at $t-1$, simple-averaged across years within each sub-period. Pre-pol = 1984–1993; Pol = pre-crisis 1994–2006; Post-crisis = 2009–2019.

B.3.2 Five-State Unconditional Transition Matrices

Tables 21 and 22 report the simple-average unconditional (joint) five-state transition matrices for the pre-polarization (1984–1993) and polarization (1994–2021) sub-periods. Each cell gives the unconditional probability that a randomly drawn working-age individual is in state κ at t and state κ' at $t+1$; the matrix sums to one. Row sums give the origin distribution (the population share in state κ at t); column sums give the destination distribution. Comparing the two panels makes the key rc-specific transitions visible as joint probabilities: the $o \to rc$ cell falls while the $rc \to rc$ diagonal rises, and the rm row compresses as the rm population itself shrinks. Four gross-flow patterns jump out when comparing the two sub-periods — patterns that the net-flow decomposition of Table 14 collapses into single signed numbers. First, the $o \to rc$ cell falls from 2.49 to 1.76 (a 29% drop), while the $rc \to o$ cell falls only from 3.02 to 2.05 (a 32% drop); the routine-cognitive sign flip documented in the main text is therefore driven primarily by the entry side, not the exit side. Second, the $o \to rm$ cell falls even more steeply, from 1.55 to 1.06 (32% down), but so does the reverse $rm \to o$ flow (2.66 to 1.39, a 48% drop), leaving the rm net flow from non-employment almost unchanged. Third, the $rc \to rc$ diagonal rises (18.55 to 20.65) even as the routine population contracts, so the rise must come from higher within-state persistence rather than from the bin — consistent with lengthening occupation-specific tenure. Fourth, the $m \to m$ and $a \to a$ diagonals also rise

Table 21: Unconditional transition matrix — pre-polarization 1984–1993. Values $\times 100$.

From κ \ To κ'	m	rc	rm	a	o	Row sum
m	7.63	1.07	1.22	0.14	1.96	12.02
rc	0.87	18.55	0.85	1.32	3.02	24.60
rm	1.20	1.06	15.98	0.38	2.66	21.28
a	0.10	1.00	0.28	6.58	1.16	9.12
o	1.32	2.49	1.55	0.83	26.79	32.97
Col sum	11.12	24.17	19.88	9.24	35.58	100.00

Table 22: Unconditional transition matrix — polarization 1994–2021. Values $\times 100$.

From κ \ To κ'	m	rc	rm	a	o	Row sum
m	10.89	0.92	0.71	0.16	1.79	14.48
rc	0.79	20.65	0.53	0.96	2.05	24.99
rm	0.72	0.59	11.60	0.28	1.39	14.58
a	0.10	0.68	0.18	10.89	0.91	12.77
o	1.60	1.76	1.06	0.78	27.99	33.19
Col sum	14.10	24.61	14.09	13.06	34.14	100.00

(7.63 to 10.89 and 6.58 to 10.89), but the manual and abstract populations expand over the sample — the a row sum rises 40% (9.12 to 12.77) — so for these states the gains track the growing bins and cannot be read as higher persistence. Column sums likewise show the polarization pattern in gross terms: the rm column shrinks from 19.88 to 14.09, m expands from 11.12 to 14.10, and a expands from 9.24 to 13.06.

B.4 Linear Probability Models: Full Regression Tables

This appendix reports the full LPM estimates behind the summary coefficients of Tables 4 and 5. For each origin state $\kappa \in \{m, r, a, o\}$ and destination $\kappa' \neq \kappa$, the dependent variable indicates the one-year transition $\kappa \to \kappa'$. Every regression includes a constant, age, a female dummy, a college indicator, tenure (years in the current two-digit PCS code for employed origins; ongoing non-employment-spell duration for o-origin), and year fixed effects, with each year weighted equally and HC3-robust standard errors. College is back-projected within worker from the first year it is observed; for o-origin transitions the sample is restricted to re-entrants. The rate $\bar{y}$ at the foot of each table is the unconditional transition probability within the origin-state sample.

Each table reports the regression on the full panel and separately on the pre-polarization (1984–1993) and polarization (1994–2021) sub-samples, so its first three columns reproduce the coefficients in Tables 4–5. For the nine employed-origin transitions, three further columns add an indicator for the origin job being in the public sector (central or local government, or public hospitals) and an indicator for a fixed-term contract; both are undefined off an origin job, so the three re-entry transitions carry the baseline columns only.

Table 23: Full LPM regressions: $r \to a$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	$+0.0587^{***}$	$+0.0680^{***}$	$+0.0421^{***}$	$+0.0597^{***}$	$+0.0685^{***}$	$+0.0435^{***}$
	(0.0015)	(0.0017)	(0.0008)	(0.0015)	(0.0017)	(0.0008)
Age	$+0.0004^{***}$	$+0.0013^{***}$	$+0.0002^{***}$	$+0.0005^{***}$	$+0.0013^{***}$	$+0.0003^{***}$
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Female	-0.0219^{***}	-0.0264^{***}	-0.0196^{***}	-0.0197^{***}	-0.0256^{***}	-0.0171^{***}
	(0.0002)	(0.0008)	(0.0002)	(0.0003)	(0.0008)	(0.0002)
College	$+0.0966^{***}$	$+0.2017^{***}$	$+0.0807^{***}$	$+0.0983^{***}$	$+0.2024^{***}$	$+0.0824^{***}$
	(0.0008)	(0.0050)	(0.0006)	(0.0008)	(0.0049)	(0.0006)
Tenure	-0.0031^{***}	-0.0065^{***}	-0.0029^{***}	-0.0031^{***}	-0.0065^{***}	-0.0029^{***}
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Public sector				-0.0131^{***}	-0.0063^{***}	-0.0139^{***}
				(0.0002)	(0.0011)	(0.0002)
Fixed-term (CDD)				-0.0094^{**}	-0.0118	-0.0067^{*}
				(0.0032)	(0.0062)	(0.0030)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	8,260,353	244,532	8,015,821	8,265,381	247,431	8,017,950
R^2	0.0266	0.0426	0.0242	0.0275	0.0430	0.0253
$\bar{y}$	0.0326	0.0405	0.0323	0.0326	0.0408	0.0323

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 24: Full LPM regressions: $m \to a$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	$+0.0172^{***}$	$+0.0187^{***}$	$+0.0147^{***}$	$+0.0168^{***}$	$+0.0183^{***}$	$+0.0142^{***}$
	(0.0017)	(0.0019)	(0.0009)	(0.0016)	(0.0019)	(0.0009)
Age	-0.0000	$+0.0001$	-0.0000^{**}	$+0.0000$	$+0.0001$	-0.0000^{**}
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Female	-0.0090^{***}	-0.0090^{***}	-0.0088^{***}	-0.0087^{***}	-0.0090^{***}	-0.0085^{***}
	(0.0003)	(0.0010)	(0.0002)	(0.0003)	(0.0010)	(0.0002)
College	$+0.0773^{***}$	$+0.1181^{***}$	$+0.0700^{***}$	$+0.0771^{***}$	$+0.1191^{***}$	$+0.0696^{***}$
	(0.0022)	(0.0116)	(0.0015)	(0.0022)	(0.0116)	(0.0015)
Tenure	-0.0013^{***}	-0.0024^{***}	-0.0012^{***}	-0.0013^{***}	-0.0023^{***}	-0.0012^{***}
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Public sector				$+0.0085^{***}$	$+0.0124^{***}$	$+0.0072^{***}$
				(0.0007)	(0.0033)	(0.0005)
Fixed-term (CDD)				-0.0061^{***}	-0.0100^{**}	-0.0058^{***}
				(0.0004)	(0.0033)	(0.0003)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2,899,460	56,363	2,843,097	2,901,503	57,431	2,844,072
R^2	0.0203	0.0228	0.0203	0.0208	0.0237	0.0207
$\bar{y}$	0.0105	0.0122	0.0105	0.0105	0.0123	0.0105

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 25: Full LPM regressions: $m \to r$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	$+0.2515^{***}$	$+0.2947^{***}$	$+0.1509^{***}$	$+0.2479^{***}$	$+0.2885^{***}$	$+0.1494^{***}$
	(0.0063)	(0.0069)	(0.0034)	(0.0061)	(0.0068)	(0.0034)
Age	-0.0012^{***}	-0.0005	-0.0013^{***}	-0.0012^{***}	-0.0008^{**}	-0.0013^{***}
	(0.0000)	(0.0003)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Female	-0.0376^{***}	-0.0645^{***}	-0.0285^{***}	-0.0359^{***}	-0.0645^{***}	-0.0263^{***}
	(0.0010)	(0.0034)	(0.0007)	(0.0010)	(0.0034)	(0.0007)
College	$+0.0574^{***}$	$+0.0302^{*}$	$+0.0608^{***}$	$+0.0554^{***}$	$+0.0265$	$+0.0592^{***}$
	(0.0029)	(0.0150)	(0.0021)	(0.0028)	(0.0149)	(0.0021)
Tenure	-0.0112^{***}	-0.0358^{***}	-0.0096^{***}	-0.0112^{***}	-0.0351^{***}	-0.0098^{***}
	(0.0001)	(0.0008)	(0.0001)	(0.0001)	(0.0008)	(0.0001)
Public sector				$+0.0359^{***}$	$+0.0409^{***}$	$+0.0310^{***}$
				(0.0019)	(0.0086)	(0.0014)
Fixed-term (CDD)				-0.0336^{***}	$+0.0081$	-0.0401^{***}
				(0.0022)	(0.0187)	(0.0014)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2,899,460	56,363	2,843,097	2,901,503	57,431	2,844,072
R^2	0.0504	0.0448	0.0363	0.0511	0.0453	0.0374
$\bar{y}$	0.0840	0.1747	0.0822	0.0840	0.1741	0.0822

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 26: Full LPM regressions: $a \to r$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	$+0.2220^{***}$	$+0.2397^{***}$	$+0.1199^{***}$	$+0.2181^{***}$	$+0.2357^{***}$	$+0.1163^{***}$
	(0.0066)	(0.0069)	(0.0029)	(0.0064)	(0.0067)	(0.0028)
Age	-0.0011^{***}	-0.0021^{***}	-0.0008^{***}	-0.0013^{***}	-0.0022^{***}	-0.0011^{***}
	(0.0000)	(0.0003)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Female	$+0.0365^{***}$	$+0.0511^{***}$	$+0.0310^{***}$	$+0.0316^{***}$	$+0.0465^{***}$	$+0.0262^{***}$
	(0.0011)	(0.0043)	(0.0007)	(0.0010)	(0.0042)	(0.0007)
College	-0.0649^{***}	-0.0982^{***}	-0.0546^{***}	-0.0702^{***}	-0.1037^{***}	-0.0598^{***}
	(0.0009)	(0.0034)	(0.0006)	(0.0009)	(0.0034)	(0.0006)
Tenure	-0.0095^{***}	-0.0278^{***}	-0.0086^{***}	-0.0092^{***}	-0.0272^{***}	-0.0084^{***}
	(0.0001)	(0.0008)	(0.0001)	(0.0001)	(0.0008)	(0.0001)
Public sector				$+0.0404^{***}$	$+0.0511^{***}$	$+0.0379^{***}$
				(0.0012)	(0.0058)	(0.0008)
Fixed-term (CDD)				$+0.0801^{*}$	$+0.1119$	$+0.0461^{*}$
				(0.0334)	(0.0644)	(0.0215)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	3,013,614	52,856	2,960,758	3,016,089	54,183	2,961,906
R^2	0.0619	0.0517	0.0507	0.0636	0.0529	0.0533
$\bar{y}$	0.0725	0.1492	0.0711	0.0725	0.1483	0.0711

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 27: Full LPM regressions: $r \to m$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	$+0.0460^{***}$	$+0.0498^{***}$	$+0.0375^{***}$	$+0.0465^{***}$	$+0.0510^{***}$	$+0.0384^{***}$
	(0.0014)	(0.0016)	(0.0008)	(0.0014)	(0.0015)	(0.0008)
Age	-0.0003^{***}	-0.0004^{***}	-0.0003^{***}	-0.0003^{***}	-0.0003^{***}	-0.0003^{***}
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Female	-0.0041^{***}	-0.0028^{***}	-0.0041^{***}	-0.0023^{***}	-0.0006	-0.0025^{***}
	(0.0002)	(0.0008)	(0.0002)	(0.0002)	(0.0008)	(0.0002)
College	-0.0150^{***}	-0.0197^{***}	-0.0145^{***}	-0.0138^{***}	-0.0192^{***}	-0.0133^{***}
	(0.0003)	(0.0019)	(0.0003)	(0.0003)	(0.0019)	(0.0003)
Tenure	-0.0037^{***}	-0.0086^{***}	-0.0034^{***}	-0.0037^{***}	-0.0085^{***}	-0.0034^{***}
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Public sector				-0.0112^{***}	-0.0184^{***}	-0.0100^{***}
				(0.0002)	(0.0009)	(0.0002)
Fixed-term (CDD)				$+0.0716^{***}$	$+0.0817^{***}$	$+0.0614^{***}$
				(0.0065)	(0.0133)	(0.0051)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	8,260,353	244,532	8,015,821	8,265,381	247,431	8,017,950
R^2	0.0108	0.0086	0.0116	0.0117	0.0101	0.0124
$\bar{y}$	0.0281	0.0359	0.0279	0.0282	0.0361	0.0279

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 28: Full LPM regressions: $a \to m$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	$+0.0154^{***}$	$+0.0155^{***}$	$+0.0101^{***}$	$+0.0156^{***}$	$+0.0157^{***}$	$+0.0097^{***}$
	(0.0019)	(0.0020)	(0.0007)	(0.0019)	(0.0020)	(0.0007)
Age	-0.0002^{***}	-0.0003^{***}	-0.0002^{***}	-0.0002^{***}	-0.0002^{***}	-0.0002^{***}
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Female	$+0.0016^{***}$	$+0.0064^{***}$	$+0.0003$	$+0.0010^{**}$	$+0.0054^{***}$	-0.0002
	(0.0003)	(0.0014)	(0.0002)	(0.0003)	(0.0013)	(0.0002)
College	-0.0075^{***}	-0.0069^{***}	-0.0076^{***}	-0.0081^{***}	-0.0079^{***}	-0.0081^{***}
	(0.0003)	(0.0010)	(0.0002)	(0.0003)	(0.0010)	(0.0002)
Tenure	-0.0011^{***}	-0.0026^{***}	-0.0010^{***}	-0.0010^{***}	-0.0025^{***}	-0.0009^{***}
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Public sector				$+0.0042^{***}$	$+0.0088^{***}$	$+0.0034^{***}$
				(0.0004)	(0.0020)	(0.0002)
Fixed-term (CDD)				$+0.0147$	$+0.0060$	$+0.0223^{***}$
				(0.0093)	(0.0182)	(0.0049)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	3,013,614	52,856	2,960,758	3,016,089	54,183	2,961,906
R^2	0.0057	0.0043	0.0065	0.0060	0.0048	0.0067
$\bar{y}$	0.0086	0.0102	0.0086	0.0086	0.0103	0.0086

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 29: Full LPM regressions: $m \to o$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	+0.1554***	+0.1345***	+0.1773***	+0.1582***	+0.1412***	+0.1802***
	(0.0054)	(0.0059)	(0.0039)	(0.0053)	(0.0058)	(0.0038)
Age	−0.0005***	−0.0029***	−0.0002***	−0.0004***	−0.0024***	−0.0001***
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Female	+0.0287***	+0.0447***	+0.0235***	+0.0283***	+0.0454***	+0.0228***
	(0.0009)	(0.0032)	(0.0007)	(0.0009)	(0.0032)	(0.0007)
College	+0.0235***	+0.0560***	+0.0179***	+0.0252***	+0.0582***	+0.0195***
	(0.0027)	(0.0143)	(0.0019)	(0.0027)	(0.0143)	(0.0019)
Tenure	−0.0130***	−0.0189***	−0.0128***	−0.0130***	−0.0197***	−0.0128***
	(0.0001)	(0.0009)	(0.0001)	(0.0001)	(0.0009)	(0.0001)
Public sector				−0.0371***	−0.0387***	−0.0366***
				(0.0016)	(0.0069)	(0.0012)
Fixed-term (CDD)				−0.0166***	−0.0926***	−0.0074***
				(0.0020)	(0.0141)	(0.0016)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2,899,460	56,363	2,843,097	2,901,503	57,431	2,844,072
R^2	0.0325	0.0388	0.0290	0.0332	0.0398	0.0296
$\bar{y}$	0.1433	0.1737	0.1427	0.1433	0.1740	0.1427

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 30: Full LPM regressions: $r \to o$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	+0.1318***	+0.1143***	+0.1230***	+0.1333***	+0.1183***	+0.1259***
	(0.0023)	(0.0025)	(0.0015)	(0.0023)	(0.0025)	(0.0015)
Age	+0.0006***	−0.0018***	+0.0009***	+0.0007***	−0.0015***	+0.0010***
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Female	+0.0038***	+0.0163***	+0.0002	+0.0081***	+0.0211***	+0.0045***
	(0.0004)	(0.0013)	(0.0003)	(0.0004)	(0.0014)	(0.0003)
College	−0.0019**	+0.0238***	−0.0048***	+0.0014*	+0.0265***	−0.0016**
	(0.0007)	(0.0042)	(0.0005)	(0.0007)	(0.0041)	(0.0005)
Tenure	−0.0082***	−0.0154***	−0.0079***	−0.0081***	−0.0154***	−0.0078***
	(0.0000)	(0.0003)	(0.0000)	(0.0000)	(0.0003)	(0.0000)
Public sector				−0.0282***	−0.0410***	−0.0257***
				(0.0004)	(0.0018)	(0.0003)
Fixed-term (CDD)				+0.0794***	+0.0581***	+0.0914***
				(0.0077)	(0.0151)	(0.0068)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	8,260,353	244,532	8,015,821	8,265,381	247,431	8,017,950
R^2	0.0219	0.0274	0.0190	0.0235	0.0290	0.0206
$\bar{y}$	0.0924	0.1176	0.0916	0.0924	0.1182	0.0916

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 31: Full LPM regressions: $a \to o$

	Baseline			+ Public, CDD		
	84–21	84–93	94–21	84–21	84–93	94–21
Constant	$+0.1286^{***}$	$+0.1264^{***}$	$+0.1006^{***}$	$+0.1281^{***}$	$+0.1277^{***}$	$+0.1015^{***}$
	(0.0052)	(0.0054)	(0.0028)	(0.0051)	(0.0053)	(0.0028)
Age	$+0.0013^{***}$	-0.0005^{*}	$+0.0015^{***}$	$+0.0013^{***}$	-0.0002	$+0.0016^{***}$
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0000)
Female	$+0.0097^{***}$	$+0.0202^{***}$	$+0.0061^{***}$	$+0.0103^{***}$	$+0.0207^{***}$	$+0.0068^{***}$
	(0.0009)	(0.0035)	(0.0006)	(0.0009)	(0.0035)	(0.0006)
College	-0.0090^{***}	-0.0071^{*}	-0.0090^{***}	-0.0086^{***}	-0.0078^{*}	-0.0081^{***}
	(0.0008)	(0.0031)	(0.0005)	(0.0008)	(0.0031)	(0.0005)
Tenure	-0.0062^{***}	-0.0176^{***}	-0.0057^{***}	-0.0063^{***}	-0.0178^{***}	-0.0058^{***}
	(0.0001)	(0.0008)	(0.0001)	(0.0001)	(0.0007)	(0.0001)
Public sector				-0.0050^{***}	-0.0014	-0.0064^{***}
				(0.0010)	(0.0048)	(0.0008)
Fixed-term (CDD)				$+0.0490$	$+0.0280$	$+0.0666^{**}$
				(0.0273)	(0.0505)	(0.0231)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	3,013,614	52,856	2,960,758	3,016,089	54,183	2,961,906
R^2	0.0189	0.0210	0.0144	0.0191	0.0209	0.0146
$\bar{y}$	0.0698	0.1124	0.0690	0.0699	0.1131	0.0691

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 32: Full LPM regressions: $o \to r$

	Baseline		
	84–21	84–93	94–21
Constant	$+0.1219^{***}$	$+0.2264^{***}$	$+0.2827^{***}$
	(0.0024)	(0.0030)	(0.0023)
Age	-0.0024^{***}	$+0.0011^{***}$	-0.0026^{***}
	(0.0000)	(0.0001)	(0.0000)
Female	-0.0113^{***}	-0.0284^{***}	-0.0034^{***}
	(0.0005)	(0.0016)	(0.0004)
College	-0.0470^{***}	-0.0776^{***}	-0.0397^{***}
	(0.0006)	(0.0020)	(0.0005)
Tenure (NE spell)	-0.0160^{***}	-0.0804^{***}	-0.0134^{***}
	(0.0001)	(0.0011)	(0.0001)
Year FE	Yes	Yes	Yes
N	6,482,355	184,634	6,297,721
R^2	0.0384	0.0540	0.0410
$\bar{y}$	0.0898	0.1353	0.0885

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 33: Full LPM regressions: $o \to m$

	Baseline		
	84–21	84–93	94–21
Constant	$+0.0268^{***}$	$+0.0580^{***}$	$+0.0856^{***}$
	(0.0014)	(0.0018)	(0.0015)
Age	-0.0014^{***}	-0.0002^{*}	-0.0014^{***}
	(0.0000)	(0.0001)	(0.0000)
Female	$+0.0148^{***}$	$+0.0021^{*}$	$+0.0196^{***}$
	(0.0003)	(0.0009)	(0.0003)
College	-0.0424^{***}	-0.0350^{***}	-0.0446^{***}
	(0.0003)	(0.0009)	(0.0003)
Tenure (NE spell)	-0.0057^{***}	-0.0213^{***}	-0.0050^{***}
	(0.0001)	(0.0007)	(0.0001)
Year FE	Yes	Yes	Yes
N	6,482,355	184,634	6,297,721
R^2	0.0181	0.0129	0.0203
$\bar{y}$	0.0569	0.0414	0.0574

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 34: Full LPM regressions: $o \to a$

	Baseline		
	84–21	84–93	94–21
Constant	$+0.0314^{***}$	$+0.0604^{***}$	$+0.0585^{***}$
	(0.0011)	(0.0016)	(0.0011)
Age	-0.0001^{***}	$+0.0013^{***}$	-0.0002^{***}
	(0.0000)	(0.0001)	(0.0000)
Female	-0.0193^{***}	-0.0258^{***}	-0.0166^{***}
	(0.0003)	(0.0008)	(0.0002)
College	$+0.0712^{***}$	$+0.0686^{***}$	$+0.0716^{***}$
	(0.0006)	(0.0022)	(0.0005)
Tenure (NE spell)	-0.0040^{***}	-0.0181^{***}	-0.0034^{***}
	(0.0000)	(0.0005)	(0.0000)
Year FE	Yes	Yes	Yes
N	6,482,355	184,634	6,297,721
R^2	0.0481	0.0340	0.0547
$\bar{y}$	0.0291	0.0270	0.0292

Notes: Coefficients in probability units, HC3 SE in parentheses; each year weighted equally. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

The augmented columns add, for the nine employed-origin transitions, an indicator for the origin job being in the public sector and one for a fixed-term contract (CDD). Fixed-term contracts are strongly associated with leaving employment and with downgrading, and their tie to exit strengthens over the period: the CDD coefficient on $r \to o$ rises from +5.8 pp pre-1994 to +9.1 during polarization (and on $a \to o$ from +2.8 to +6.7), consistent with fixed-term separation becoming a more important route into non-employment. Public-sector origin runs the other way on that margin — it lowers exits ($m \to o$, -3.7; $r \to o$, -2.8), the familiar attachment-stabilizing role — while on the upgrading transitions its sign is mixed ($r \to a$, -1.3; $m \to r$, $+3.6$). Adding the two controls leaves the gender, college, and tenure coefficients of the baseline columns broadly unchanged.

B.4.1 Tenure and the Panel Length

Table 35 re-estimates the period split with tenure (and, for re-entry, non-employment-spell duration) capped at nine years in both sub-samples, the maximum tenure observable in the pre-polarization panel. Because tenure is otherwise bounded by panel length — at most nine years before 1994 against roughly thirty-seven during polarization — the apparent post-1994 weakening of tenure as a stabilizer (Section 5.2.2) could be a mechanical artefact of the longer right tail. Capping addresses this directly. For the employed-origin transitions the average polarization-to-pre-polarization tenure-coefficient ratio rises from about 0.4 uncapped to 0.7 capped, so most of the apparent weakening is mechanical. On the re-entry margin the ratio is about 0.2 and barely moves under capping, so there the post-1994 attenuation of the non-employment-duration coefficient is genuine.

C Robustness

C.1 30–59 Age Window

The 25–64 age window creates demographic inflows (aging-in at 25) and outflows (aging-out at 64) that are absorbed by the non-employment channel. To assess whether these boundary flows affect the results, we replicate the full analysis on the narrower 30–59 age window, which excludes most first-time labor-market entries (concentrated at ages 25–29) and the retirement transitions concentrated around the legal early-retirement age.

The bilateral JtJ flows are virtually identical across the two age windows (differences < 0.01 pp/year for every bilateral flow and every sub-period). This is expected: boundary workers have from_state $= o$ or to_state $= o$ by construction and therefore never appear in JtJ flows.

The manual Excess_NE drops by about 40% (from +0.19 to +0.11 pp/year), reflecting the removal of the manual-overweight aging-in flow at 25. The abstract Excess_NE moves from -0.19 to -0.13 pp/year, a +0.06 pp/year improvement (about 30%), again

Table 35: Period split with tenure capped at nine years (cf. Table 5)

Transition	**Pre-polarization (1984–93)** College	Female	Tenure	**Polarization (1994–2021)** College	Female	Tenure
Within-employment upgrading						
$m \to r$	$+0.030^{*}$	-0.065^{***}	-0.0358^{***}	$+0.058^{***}$	-0.028^{***}	-0.0151^{***}
$r \to a$	$+0.202^{***}$	-0.026^{***}	-0.0065^{***}	$+0.080^{***}$	-0.019^{***}	-0.0047^{***}
$m \to a$	$+0.118^{***}$	-0.009^{***}	-0.0024^{***}	$+0.070^{***}$	-0.009^{***}	-0.0019^{***}
Within-employment downgrading						
$r \to m$	-0.020^{***}	-0.003^{***}	-0.0086^{***}	-0.015^{***}	-0.004^{***}	-0.0058^{***}
$a \to r$	-0.098^{***}	$+0.051^{***}$	-0.0278^{***}	-0.053^{***}	$+0.031^{***}$	-0.0138^{***}
$a \to m$	-0.007^{***}	$+0.006^{***}$	-0.0026^{***}	-0.007^{***}	$+0.000$	-0.0016^{***}
Exits to non-employment						
$m \to o$	$+0.056^{***}$	$+0.045^{***}$	-0.0189^{***}	$+0.014^{***}$	$+0.024^{***}$	-0.0194^{***}
$r \to o$	$+0.024^{***}$	$+0.016^{***}$	-0.0154^{***}	-0.006^{***}	$+0.001^{***}$	-0.0130^{***}
$a \to o$	-0.007^{*}	$+0.020^{***}$	-0.0176^{***}	-0.008^{***}	$+0.006^{***}$	-0.0087^{***}
(Re-)entries from non-employment						
$o \to m$	-0.035^{***}	$+0.002^{*}$	-0.0213^{***}	-0.045^{***}	$+0.020^{***}$	-0.0060^{***}
$o \to r$	-0.078^{***}	-0.028^{***}	-0.0804^{***}	-0.039^{***}	-0.003^{***}	-0.0162^{***}
$o \to a$	$+0.069^{***}$	-0.026^{***}	-0.0181^{***}	$+0.072^{***}$	-0.017^{***}	-0.0041^{***}

Notes: As Table 5 but with occupation tenure (and, for re-entry, non-employment-spell duration) capped at nine years in both sub-periods, the pre-polarization panel-length maximum. Each year weighted equally; HC3 SE. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

consistent with the retirement-cliff exits being disproportionately drawn from abstract employment. Crucially, the routine Excess_NE sign flip survives the narrower age window and is in fact slightly larger: under 30–59, routine Excess_NE moves from +0.31 pp/year (pre-polarization) to −0.08 pp/year (polarization), a swing of −0.39 pp/year, compared with −0.34 pp/year under 25–64.

C.2 Alternative Identification of Non-Employment Entries

Section 3.2 contrasts two ways of identifying the gross entries from non-employment, $F_{o,k}$: the baseline *identity-based recovery* and the alternative *direct tabulation*. This appendix develops the comparison and shows that the two are equivalent for every quantity that enters the employment-share decomposition, differing only in a cell that does not.

The decomposition of ΔE_k depends only on the *net* non-employment flow $F_{o,k} - F_{k,o}$, which the accounting identity (6) pins down exactly. The transition *matrix*, by contrast, requires the *gross* entry flow $F_{o,k}$ on its own, and because the 25–64 age restriction is applied independently at $t-1$ and t, this gross flow pools three distinct populations: (A) genuine non-employment-to-employment moves; (B) cohort aging-in (workers who turn 25 between $t-1$ and t); and (C) panel first-entries (workers first observed in DADS–EDP at t). The two methods differ in how they construct this flow.

Identity-based recovery (the baseline) observes job-to-job flows and exits $F_{k,o}$ from the age-boundary-augmented matched panel and recovers the entries as the residual $F_{o,k} = \Delta N_k - \text{Net\_JtJ}_k + F_{k,o}$, setting the $o \to o$ cell from the INSEE non-employment stock. It requires no coding of pre-entry non-employment. *Direct tabulation* instead constructs the full 4×4 matrix by cross-tabulating each worker's origin and destination state in the same augmented panel, reading $F_{o,k}$, $F_{k,o}$ and $F_{o,o}$ off the data directly; this requires coding each worker's pre-entry years as genuine non-employment, except for the structural gaps created by the 2002 and 2006 sample expansions.[19]

With the same age-boundary imputation applied to both, direct tabulation exactly reproduces identity-based recovery: the entries $F_{o,k}$, exits $F_{k,o}$, job-to-job flows, and the resulting ΔE_k are identical, and both satisfy the accounting identity with zero residual. This equality is forced by the identity rather than coincidental: the two methods share the job-to-job flows and the exits $F_{k,o}$ (employed-origin flows, identical under either pre-entry coding), so once the directly-tabulated matrix balances at zero residual the identity leaves a single admissible value for its entries $F_{o,k}$, destination by destination. The lone exceptionis the 1989→1991 gap pair, whose two-year flows are halved to annual equivalents: there the age-boundary imputation—calibrated for one-year transitions—leaves a small discrepancy

[19]For a worker with panel-entry year E, the origin state at year t (within the 25–64 window) is set to non-employment whenever $E > t$, except: missing if $E \geq 2002$ and $t < 2002$ (the 2002-expansion structural gap) or $E = 2006$ and $t < 2006$ (the 2006 expansion). Both methods additionally apply the same age-boundary imputation, coding boundary-crossers as $o \to k$ entries (aging-in at 25) or $k \to o$ exits (aging-out at 64).

in the recovered entries (the routine Excess_NE channel moves by at most 0.03 pp). This single pair does not affect the period averages.

The methods diverge only in the $o \to o$ cell. Table 36 reports the two simple-average conditional matrices: the three employed rows are identical, and only the non-employment row differs. Direct tabulation counts $F_{o,o}$ from the panel—every worker non-employed in both years—yielding a higher non-employment persistence (0.81 vs. 0.69), because the DADS panel non-employment pool exceeds the INSEE-implied one. This larger $o \to o$ mass merely rescales the conditional $o \to k$ probabilities in that row; the underlying *gross* entries $F_{o,k}$ are the same. The difference therefore affects conditional non-employment-row probabilities but not the employment-share decomposition, which depends solely on $F_{o,k}$, $F_{k,o}$, and job-to-job flows.

Table 36: Simple-average conditional transition matrices, pooled over 1984–2021: direct tabulation vs. identity-based recovery. Rows sum to one. The three employed rows are identical across the two methods; only the non-employment (o) row differs.

From κ \ To κ'	o	m	r	a
Direct tabulation				
o	0.808	0.053	0.112	0.027
m	0.161	0.721	0.108	0.010
r	0.107	0.030	0.830	0.033
a	0.093	0.008	0.085	0.814
Identity-based recovery (baseline)				
o	0.690	0.085	0.182	0.043
m	0.161	0.721	0.108	0.010
r	0.107	0.030	0.830	0.033
a	0.093	0.008	0.085	0.814

The equivalence confirms that the entry-margin collapse documented in Section 4—a property of the recovered $F_{o,k}$—is not an artefact of the identity-based construction: a direct count of non-employment-to-employment transitions delivers the same flows.

C.3 The Role of Self-Employment in the o State

The state o in our framework pools genuine non-employment with self-employment and other DADS-uncovered states. We address two distinct (but interacting) concerns. First, general self-employment (SE) churn contaminates the $o \to k$ and $k \to o$ flows (§C.3.1). Second, the 2009 DADS-grand-format expansion — the auto-entrepreneur regime plus the CESU integration — produces a stock jump in CS2=56 (manual) in 2009 (next subsection). Their interaction is discussed at the end.

C.3.1 Controlling for self-employment

A. Setup and notation. The state o pools genuine non-employment (o_1) with self-employment (o_2). DADS cannot distinguish the two sides of o. Both the observed $k \to o$ exits and the recovered $o \to k$ entries (via the accounting identity (6)) therefore pool genuine NE flows with SE-mediated ones. In other words, some apparent NE entries into occupation k are actually ex-self-employed workers entering salaried k, and some apparent NE exits from occupation k are actually salaried workers becoming self-employed. This subsection assesses whether this pooling affects the headline results of §4.

Equation 8 from section 3 can be written as

$$\Delta E_k(t) \;=\; \frac{\underbrace{\sum_{k' \in \{m,r,a\},\, k' \neq k} \Big[F_{k',k}(t) - F_{k,k'}(t)\Big]}_{\text{Net\_JtJ}_k(t)} \;+\; \underbrace{F_{o,k}(t) - F_{k,o}(t) - E_k(t) \cdot \Delta N_e(t)}_{\text{Excess\_NE}_k(t)}}{N_e(t{-}1)}, \tag{11}$$

where $F_{\kappa,\kappa'}(t)$ stands for the gross flow of workers from state κ at $t-1$ to state κ' at t. We first observe that the *job-to-job channel is self-employment-free by construction.* Its numerator counts only direct salaried-to-salaried transitions. A move $k' \to o_2 \to k$ over consecutive years is carried by the Excess_NE$_k$ channel, not by job-to-job. So Net_JtJ$_k$ is not contaminated by self-employment flows.

By contrast, the observed net flow from non-employment $F_{o,k} - F_{k,o}$ pools both genuine non-employment (o_1) and self-employment (o_2). Define

$$\nu_k(t) \;\equiv\; F_{o_2,k}(t) - F_{k,o_2}(t), \qquad \nu(t) \;\equiv\; \sum_{k \in \{m,r,a\}} \nu_k(t),$$

the (unobserved) net annual flow from self-employment into salaried occupation k — ex-self-employed workers entering salaried-k minus salaried-k workers becoming self-employed — and its aggregate. Analogously, we define

$$\eta_k(t) \;\equiv\; F_{o_1,k}(t) - F_{k,o_1}(t), \qquad \eta(t) \;\equiv\; \sum_{k \in \{m,r,a\}} \eta_k(t),$$

the (unobserved) genuine net non-employment flow from and to occupation k and its aggregate. Hence we have $\eta_k = (F_{o,k} - F_{k,o}) - \nu_k$. The salaried stock changes only through the o-channel — job-to-job merely reshuffles workers within salaried employment and sums to zero — so $\Delta N_e(t) = \sum_k [F_{o,k}(t) - F_{k,o}(t)] = \eta(t) + \nu(t)$. We can then split the Excess_NE$_k$ channel as follows:

$$\frac{\text{Excess\_NE}_k(t)}{N_e(t-1)} \;=\; \underbrace{\frac{\eta_k(t) - E_k(t)\,\eta(t)}{N_e(t-1)}}_{\text{genuine non-employment}} \;+\; \underbrace{\frac{\nu_k(t) - E_k(t)\,\nu(t)}{N_e(t-1)}}_{\text{self-employment}} \tag{12}$$

Our aim is therefore to provide a quantification for the unbiased

$$\frac{\text{Excess\_NE}_k^{\text{NE-only}}(t)}{N_e(t-1)} = \frac{\text{Excess\_NE}_k(t)}{N_e(t-1)} - \Xi_k(t) \tag{13}$$

where $\Xi_k = \frac{\nu_k(t) - E_k(t)\,\nu(t)}{N_e(t-1)}$, defined as the difference between the observed and the corrected one, is the *self-employment bias* in the observed Excess_NE$_k$ for salaried workers.

B. Identifying ν_k. The aggregate self-employment net-flow $\nu(t) = \sum_k \nu_k(t)$ can be directly anchored in external data — through the year-on-year change in the self-employment headcount (see Section C). However, the cross-occupation *split* of this aggregate – the per-occupation net self-employment flow $\nu_k(t)$ – is not observed and must be assigned by an identifying assumption. We here set three identification assumptions and the data inputs each requires; calibrated values and sources are deferred to Section C, and the resulting bias magnitudes to Section D.

Identifications (A) and (B) read the net flow ν_k directly off the self-employment stock and need only that series; (C) instead models its gross entry and exit components separately and it is therefore the more structured and data-demanding.

To start with define

$$E_k^{\text{self}}(t) \;\equiv\; \frac{N_k^{\text{self}}(t)}{N_e^{\text{self}}(t)}$$

as the within-self-employment occupational composition, with N_k^{self} the stock of self-employed in occupation k and $N_e^{\text{self}} = \sum_k N_k^{\text{self}}$.

(A) *Proportional.* $\nu_k(t) = E_k^{\text{self}}(t)\,\nu(t)$. The cross-occupation distribution of the net SE-to-salaried flow is assumed to match the SE *stock* composition. Hence, an occupation holding, say, 30% of the self-employment stock is assigned 30% of the net flow ν, irrespective of how fast its own self-employment is actually changing. Keying on the stock *level* rather than its change, the rule over-states occupations with a large but stable self-employment stock and under-states those whose self-employment is contracting fastest.

(B) *No-switching.* $\nu_k(t) = -\Delta N_k^{\text{self}}(t)$. Workers do not change occupation when they move between self-employment and salaried status: an ex-self-employed abstract worker becomes a salaried abstract worker. Each occupation's net flow is then read directly off the change in its own SE stock, and these sum automatically to $-\Delta N_e^{\text{self}} = \nu$. Method B and method A coincide if and only if the self-employment composition is constant over time.

(C) *Occupation-specific entry and exit.* Both gross flows are conditioned on the observed occupational structure. On the *exit* side, salaried-k workers move into self-employment at an occupation-specific rate $\tau_k = (E_k^{\text{self}}/E_k)\,\tau$, where τ is the economy-wide SE-creation hazard (disciplined in Section C); this scales the hazard with k's representation in the self-employment stock, so that

$$F_{k,o_2}(t) = \tau_k(t)\,N_k(t) = E_k^{\text{self}}(t)\,\tau(t)\,N_e(t).$$

Gross SE-creation is thereby distributed across occupations in proportion to the self-employment composition E_k^{self} rather than the salaried composition E_k: occupations under-represented in self-employment relative to their salaried weight — most sharply routine — contribute proportionately little to SE creation, while over-represented occupations (manual and abstract) contribute more. On the *entry* side, ex-self-employed re-entrants are allocated across salaried occupations by the within-employment conditional transition matrix $T_{k,k'}^{\text{within}}$ estimated from DADS salaried-to-salaried flows (Tables 8 and 9 of the main text, row-normalised after dropping the o-column; distinct between the pre-1994 regime T^{pre} and the polarization regime T^{pol}):

$$F_{o_2,k'}(t) = \rho(t)\,\sum_k N_k^{\text{self}}(t)\,T_{k,k'}^{\text{within}}(t).$$

where $\rho(t)$ is the economy-wide entry rate of self-employed workers into salaried workers. This entry rate is pinned by $\nu = \sum_k \nu_k$, which — using $\sum_k \tau_k N_k = \tau N_e$ (immediate from the exit rule, since $\sum_k E_k^{\text{self}} = 1$) — gives

$$\rho(t) = \frac{\nu(t) + \tau(t)\,N_e(t)}{N_e^{\text{self}}(t)}.$$

C. Calibration. We now calibrate the inputs listed in Section B. Three external (non-DADS) series are required: the aggregate self-employment headcount $N_e^{\text{self}}(t)$, whose year-on-year change anchors ν; the within-self-employment occupational composition $E_k^{\text{self}}(t)$; and — for identification (C) — the economy-wide SE-creation hazard $\tau(t)$. Everything else (N_e, N_k, E_k, T^{within} and the occupation-specific rates τ_k) comes from DADS or follows algebraically. Values are reported as averages over the two sub-periods used throughout (pre- and post-1994).

Aggregate self-employment headcount $N_e^{self}(t)$ from INSEE *Estimations d'emploi* 1970–2018 (https://www.insee.fr/fr/statistiques/4470794). We use the year-on-year change $-\Delta N_e^{\text{self}}(t)$ as our empirical proxy for $\nu(t)$. This is an *upper bound*: workers leaving self-employment in a given year have three possible destinations — (i) salaried employment in DADS, which is exactly ν; (ii) genuine non-employment o_1 (retirement, unemployment, leaving the labour force); (iii) outside the working-age population (or death, or emigration). Only (i) shows up anywhere in our framework; (ii) and (iii) reduce N_e^{self} without producing any DADS transition. $-\Delta N_e^{\text{self}}$ therefore overstates the true ν

by the net outflows into (ii) and (iii), and our bias bounds are conservative in the same direction. Period averages are $\nu^{\text{pre}} = +81$ k/yr (1984–93, SE stock declining through the late 1980s) and $\nu^{\text{pol}} = -15$ k/yr (post-1994, as the post-2009 stock rebound comes to dominate the earlier outflow). Corresponding N_e^{self} averages, used by identifications (A) and (C), are 2.97M (1985–94) and 2.47M (1995–2018). Identification (B), however, targets the self-employment stock *change by occupation* ($\nu_k = -\Delta N_k^{\text{self}}$), for which the relevant inputs are the period *endpoints*, not these averages: the stock levels $N_e^{\text{self}} = 3.33$M (1984), 2.52M (1994) and 2.88M (2018), combined with the endpoint compositions of Table 38 (its 1984, 1994 and 2018 rows).

Within-self-employment occupational distribution $E_k^{self}(t)$. We build E_k^{self} on the occupational axis, reading it directly from labour-force-survey data where they exist and back-projecting the rest. For the polarization period we take it from Eurostat's annual cross-tabulation of self-employment by occupation (`lfsa_esgais`: self-employed persons by ISCO-08 group, France, aged 25–64), mapped to $\{m, r, a\}$ as in Table 37(A). Averaging the clean window 2015–2018 (to step over the 2011 ISCO-88→08 recoding) gives $(E_m^{\text{self}}, E_r^{\text{self}}, E_a^{\text{self}})^{\text{pol}} = (33.5, 19.9, 46.6)\%$. No occupation breakdown exists for France before 1994, so the pre-period composition is back-projected from the INSEE sectoral self-employment series via the sector-to-$\{m, r, a\}$ map of Table 37(B): routine is held at its directly-measured 20% — the sectors it maps to, industry and construction, hold a flat share of self-employment throughout the ERI20 series, pre-1994 included — so only the agriculture-to-tertiary shift moves the manual/abstract split (more agriculture, less business services in the 1980s). This yields $(\cdot)^{\text{pre}}(45.4, 19.9, 34.7)\%$, a more manual-intensive mix. The series is reported in Table 38.

Table 37: Mapping of occupations and activity sectors to $\{m, r, a\}$.

A. Eurostat ISCO-08 (`lfsa_esgais`; self-employment level, polarization)	
m	OC5 service & sales, OC6 skilled agricultural, OC9 elementary
r	OC4 clerical, OC7 craft & trades, OC8 plant/machine operators
a	OC1 managers, OC2 professionals, OC3 technicians
B. INSEE activity sector (pre-1994 back-projection)	
m	agriculture; tertiary (retail, personal services)
r	industry, construction
a	tertiary (professional & business services)

Table 38: Self-employment occupational composition E_k^{self} (in percent). Polarization period: direct (Eurostat `lfsa_esgais`, France 25–64). Pre-1994: back-projected (INSEE sectoral series). Identifications (A)/(C) use the period-average rows; (B), reading the stock change, uses the endpoint rows.

	E_m^{self}	E_r^{self}	E_a^{self}
Pre-1994 (back-projected)			
1984	48.9	19.9	31.2
1994	42.6	19.9	37.5
1985–94 average	45.4	19.9	34.7
Polarization (direct, `lfsa_esgais`)			
2015	34.0	19.9	46.1
2016	34.5	19.8	45.7
2017	33.6	20.0	46.4
2018	31.9	20.0	48.1
2015–18 average	33.5	19.9	46.6

Annual SE creation rate $\tau(t)$ from Bauer, Garbinti, and Georges-Kot (*INSEE Références* 2018, https://www.insee.fr/fr/statistiques/3641409), Figure 8 Panel B, based on the French Labour Force Survey (LFS, INSEE *Enquête Emploi*) longitudinal panel for men aged 20–64 (denominator pools *salariés* and *chômeurs*). Period averages: $\tau^{\text{pre}} = 1.4\%$ (1985–94) and $\tau^{\text{pol}} = 0.8\%$ (post-1994), the latter anchored on the 1995–2004 estimate and held across the rest of the window. An independent INSEE source corroborates this: "Qui s'installe à son compte ?" (INSEE, *Emploi et revenus des indépendants*, édition 2020; underlying figures in `data/_external/ERI20-VE.xlsx`), based on the all-active longitudinal panel over 2008–2015, reports an aggregate salarié-to-self-employment rate of about 0.7%/yr — just below the BGK 0.8% — confirming the rate's stability into the later years.[20]

The occupation-specific rates of identification (C), $\tau_k = (E_k^{\text{self}}/E_k)\,\tau$, follow from the calibrated E_k^{self}, E_k and τ, re-evaluated in each sub-period (Table 39): routine, under-represented in self-employment, creates it at about a third of the economy-wide rate; manual and abstract, over-represented, at roughly twice.

[20] The BGK figure departs from the ideal salarié-only, sex-pooled rate: men's rate is roughly twice women's, and the chômeur denominator inflates it because of the ACCRE/ARCE SE-creation subsidies, biasing it upward by at most a few hundredths of a pp/year. The ERI20 rate is the same object under a near-ideal specification (salariés-only denominator, sexes pooled), so it should sit slightly below the BGK figure — consistent with $0.7\% < 0.8\%$. We therefore keep $\tau = 0.8\%$ rather than re-estimating.

Table 39: Occupation-specific SE-creation rates $\tau_k = (E_k^{\text{self}}/E_k)\,\tau$, by sub-period.

	E_k^{self}/E_k (pre)	E_k^{self}/E_k (pol)	τ_k (pre)	τ_k (pol)
routine	0.29	0.34	0.41%	0.27%
manual	2.75	1.68	3.85%	1.34%
abstract	2.24	2.21	3.13%	1.77%

Derived quantities. The within-employment share $E_k(t)$ is taken directly from DADS (period averages of the same series plotted in Figure 2): $(E_m, E_r, E_a)^{\text{pre}} = (0.17, 0.68, 0.15)$ and $(\cdot)^{\text{pol}} = (0.20, 0.59, 0.21)$.

D. Numerical results. Applying the calibrated anchors of Section C under the three identifications of Section B gives the corrected Excess_NE$_k$ swings of Table 40.

Table 40: Corrected Excess_NE$_k$ swings (pp/year), 1984–1993 vs. 1994–2021, per-DADS-salaried units, under the three identifications of Section B.

k	observed	(A) Proportional	(B) No-switching	**(C) Occ.-specific**
m	+0.21	+0.34	+0.40	**+0.20**
r	−0.34	−0.57	−0.57	**−0.37**
a	+0.13	+0.22	+0.17	**+0.16**

Under the three identifications, the main results of Section 4 are not only preserved but reinforced. Identification (C) is the most conservative: it delivers corrected routine and abstract swings of −0.37 and 0.16 respectively, both still beyond the observed −0.34 and 0.13 in absolute value. The manual swing, by contrast, is essentially unchanged (0.20 vs. 0.21). The other two identifications, (A) and (B), are even more favourable, each returning a routine swing of −0.57 (about 1.7× the observed), an abstract swing of 0.22 and 0.17 respectively, and a manual swing of 0.34 and 0.40 respectively. In what follows we focus on (C), both as a *conservative floor* and as the most structured of the three.

The across-1994 swings in the manual and routine Excess_NE channels are the main drivers of the observed polarization break, so a legitimate concern is how sensitive the reassuring results of Table 40 are to the estimated self-employment composition. Table 41 reports this sensitivity under identification (C): it maps different pre- and post-1994 compositions into the corrected Excess_NE$_k$ swing ($k = m, r, a$), the calibrated composition included. The bias stays small — and often points in a direction that *reinforces* the paper's results — under any plausible composition. Only implausibly extreme swings move the corrected channels appreciably, and even the most adversarial composition in the entire simplex leaves the baseline intact. Because the bias is linear in E^{self}, the corrected swing attains its extremum at a *vertex* of the simplex; over all vertices, with the pre- and post-1994 compositions chosen freely and independently, the single worst point is all self-employment manual before 1994 and all routine after. Even at that global worst case the routine swing only shrinks to −0.17 and the manual to +0.12 — 50% and 57% of the observed −0.34 and +0.21 — and no composition anywhere in the simplex erodes them further.

Table 41: Sensitivity of the corrected Excess_NE$_k$ swing (identification (C), pp/year) to the assumed self-employment composition. Each row fixes a pre- and post-1994 composition $E^{\text{self}} = (m, r, a)$; the calibrated row reproduces the (C) column of Table 40.

	E^{self} pre	E^{self} post	m	r	a
calibrated	$(.45, .20, .35)$	$(.34, .20, .46)$	+0.20	−0.37	+0.16
	$(.45, .20, .35)$	$(.40, .14, .46)$	+0.21	−0.38	+0.16
	$(.40, .25, .35)$	$(.30, .20, .50)$	+0.20	−0.38	+0.17
	$(.50, .20, .30)$	$(.38, .20, .42)$	+0.21	−0.37	+0.15
	$(.40, .20, .40)$	$(.30, .20, .50)$	+0.20	−0.37	+0.17
	$(.45, .25, .30)$	$(.34, .25, .41)$	+0.21	−0.36	+0.15
	$(.42, .18, .40)$	$(.32, .16, .52)$	+0.20	−0.38	+0.17
all routine	$(0, 1, 0)$	$(0, 1, 0)$	+0.18	−0.29	+0.10
manual → routine	$(\mathbf{1}, \mathbf{0}, \mathbf{0})$	$(\mathbf{0}, \mathbf{1}, \mathbf{0})$	**+0.12**	**−0.17**	**+0.04**
abstract → routine	$(0, 0, 1)$	$(0, 1, 0)$	+0.13	−0.24	+0.11
all manual	$(1, 0, 0)$	$(1, 0, 0)$	+0.30	−0.37	+0.06

All in all, the analysis points to a limited self-employment contamination of the non-employment channel, confirming the robustness of the main results of Section 4.

C.3.2 The 2009 DADS Grand Format integration: bridging the (2008, 2010) window

Beyond general self-employment churn, a distinct concern is the 2009 DADS Grand Format (DADS-GF) measurement event. The 2009 CS2 stocks are affected by the *DADS Grand Format* (DADS-GF) integration, which combined the new auto-entrepreneur regime (effective Jan 2009, with parallel integration of CESU declarations) with more comprehensive coverage of public-sector employment. To bound the effect of this measurement event on the 1994 sign-flip finding we treat 2009 as effectively missing — analogous to how 1990 (genuinely absent from DADS) is handled throughout the paper — by replacing the consecutive pairs (2008, 2009) and (2009, 2010) with a single (2008, 2010) bridge whose flows are halved over the 2-year window.

Table 42: Excess_NE$_k$ (pp/yr). Bridged = baseline minus $\{(2008, 2009), (2009, 2010)\}$ plus $\{(2008, 2010)\}$ with halved 2-year flows.

Sub-period	ExNE$_m$	ExNE$_r$	ExNE$_a$
Pre-polarization 1984–1993	+0.04	+0.25	−0.29
Polarization 1994–2021 (baseline)	+0.25	−0.09	−0.16
(2008, 2010)-bridged	+0.20	−0.05	−0.15
1994 swing (post − pre):			
baseline	+0.21	−0.34	+0.13
(2008, 2010)-bridged	+0.16	−0.30	+0.14

The bridging has only a marginal effect on every channel. Routine Excess_NE moves from −0.09 unbridged to −0.05 bridged, with the 1994 swing attenuating from −0.34 to −0.30 pp/yr. The bilateral JtJ flows are similarly insensitive (Table 43): Net JtJ$_r$ moves from −0.44 unbridged to −0.40 bridged, and the bilateral $r \leftrightarrow a$ drain is essentially constant across 1994 in both samples (−0.54 pre-polarization, −0.52 unbridged pol, −0.51 bridged pol). The routine-to-abstract upgrading drain is structurally unchanged across the 1994 break regardless of how the crisis years are handled.

Table 43: Bilateral net JtJ flows by sub-period (pp/year). Each cell reports $(F_{k',k} - F_{k,k'})/N_e(t-1) \times 100$. Net JtJ$_k$ is the row sum.

	Bilateral net inflow into k from			
$k \setminus$ source	m	r	a	Net JtJ$_k$
Pre-polarization 1984–1993:				
manual	—	−0.20	−0.03	−0.24
routine	+0.20	—	−0.54	−0.34
abstract	+0.03	+0.54	—	+0.58
Polarization 1994–2021 (baseline):				
manual	—	−0.08	−0.05	−0.13
routine	+0.08	—	−0.52	−0.44
abstract	+0.05	+0.52	—	+0.57
Polarization 1994–2021 ((2008, 2010)-bridged):				
manual	—	−0.11	−0.05	−0.16
routine	+0.11	—	−0.51	−0.40
abstract	+0.05	+0.51	—	+0.56

The two concerns are not fully independent. General self-employment churn — the contamination of the $o \to k$ and $k \to o$ flows corrected above — is itself attenuated by the 2009 event: the auto-entrepreneur regime moves part of o_2 into DADS-salaried, shrinking the self-employment content of o, already reflected in the post-2009 self-employment rebound that drives ν^{pol} negative (Section C). The (2008, 2010) bridge then isolates the residual measurement-event component. Both effects attenuate contamination on the bridged sample, so accounting for them *strengthens* the joint robustness conclusion.

C.4 Robustness to CS2 Imputation

The DADS–EDP panel contains a sub-population of workers whose salaried wage records are observed but whose two-digit CS2 occupation code is missing in the source data. These are predominantly civil servants of the *fonction publique d'État* (FPE): the FPE administrative classification did not flow into DADS in a fully merged form before the *DADS Grand Format* (DADS-GF) integration, finalised for reference year 2009 onward. Pre-2009, DADS records the wage but not the CS2; from 2009 the CS2 codes are populated from the merged sources. Across our 1984–2021 panel, 156,389 worker-year observations (touching 126,520 unique workers, or 6.7% of the panel) fall in this category. Just over half of the volume (89,327 worker-years) sits in 2008, during the DADS-GF integration window when the legacy sample carries its most incomplete declarations; from 2009 onward the count is well under 1,000 per year.

Since the DADS–EDP panel is constructed to be representative of the French salaried workforce, dropping these worker-years would selectively remove public-sector employees from the sample and deliver a private-sector-biased panel. We therefore recover the missing CS2 codes through the imputation procedure described in Section 2. Layer 1 carries the worker's own observed CS2 from another panel year (the temporally nearest one), forward or backward; Layer 2 applies a gradient-boosted classifier trained on (log wage, sex, age, contract, sector, region, diploma, year) and predicting CS2 over the 26 valid codes; Layer 3 sets the imputed cell back to missing whenever the classifier's maximum posterior probability is below 0.30. Validation on a held-out sample yields a 4-state (manual/routine/abstract) accuracy of 89.96% and a 5-state (routine cognitive vs. routine manual) accuracy of 87.98%.

In production, Layer 1 alone fills 152,939 of the 156,389 target worker-years (97.8%) using the same worker's history; the remaining 3,450 rows (2.2%) are handled by Layer 2 or fall below the confidence threshold (in which case the cell is set to missing for that year, while the worker continues to appear in the panel for all other years where the CS2 code is observed or successfully imputed). Table 44 reports the distribution of the temporal distance between the imputed year and the source year for Layer 1 fills: 98.3% of fills draw on a source one year away from the imputed year, and 99.7% within three years. The procedure is therefore overwhelmingly a within-worker shift of one year.

To document that the headline decomposition does not depend on the imputation, we report the same exercise on a no-imputation sample: $k = m, r, a$ is set to missing for every worker-year with $CS2 = 0$ and positive earnings, and the worker-year contributes neither to stocks nor to any flow pair involving that year. The worker still appears in the panel for all other years where CS2 is observed. The intervention is more conservative than imputation, which fills these cells. Table 45 reports the main bilateral decomposition (Table 2 in the main text) under both samples.

Table 44: Distribution of temporal distance between imputed year and source year, Layer 1 worker-history fill.

Distance (years)	Imputed worker-years	Share (%)	Cumulative (%)
1	150,347	98.31	98.31
2	1,780	1.16	99.47
3	381	0.25	99.72
4	168	0.11	99.83
5	73	0.05	99.88
≥ 6	190	0.12	100.00
Total Layer 1 fills	152,939	100.00	—

Notes: For each worker-year that the Layer 1 rule fills within the imputation target set, the table reports the temporal distance $|t - t^*|$ between the imputed year t and the source year t^*, which is the closest year (forward or backward) in which the same worker's CS2 is observed. The target set is the 156,389 worker-years (126,520 unique workers) with raw CS2 missing and positive earnings; Layer 1 resolves 152,939 of them (97.8%) using the worker's own history.

Table 45: Mean annual bilateral decomposition with CS2 imputation applied versus dropping the missing-CS2 worker-years (pp/year).

Period	**Sample**		Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
Full 1984–2021	Imputed	manual	—	−0.11	−0.04	−0.16	+0.19	+0.03
		routine	+0.11	—	−0.53	−0.41	+0.00	−0.41
		abstract	+0.04	+0.53	—	+0.57	−0.19	+0.37
	No imputation	manual	—	−0.12	−0.04	−0.16	+0.20	+0.03
		routine	+0.12	—	−0.53	−0.41	+0.00	−0.40
		abstract	+0.04	+0.53	—	+0.57	−0.20	+0.37
Pre-pol 1984–1993	Imputed	manual	—	−0.20	−0.03	−0.24	+0.04	−0.20
		routine	+0.20	—	−0.54	−0.34	+0.25	−0.09
		abstract	+0.03	+0.54	—	+0.58	−0.29	+0.29
	No imputation	manual	—	−0.22	−0.04	−0.26	+0.06	−0.20
		routine	+0.22	—	−0.54	−0.32	+0.22	−0.09
		abstract	+0.04	+0.54	—	+0.58	−0.28	+0.30
Pol 1994–2021	Imputed	manual	—	−0.08	−0.05	−0.13	+0.25	+0.12
		routine	+0.08	—	−0.52	−0.44	−0.09	−0.52
		abstract	+0.05	+0.52	—	+0.57	−0.16	+0.40
	No imputation	manual	—	−0.08	−0.05	−0.13	+0.24	+0.11
		routine	+0.08	—	−0.52	−0.44	−0.08	−0.52
		abstract	+0.05	+0.52	—	+0.57	−0.17	+0.40

Notes: Values ×100 (pp/year). "Imputed" is the sample used in the main text (Table 2), in which worker-years with missing raw CS2 are filled by the procedure of Section 2. "No imputation" instead leaves these worker-years out: the occupational category is set to missing whenever the raw CS2 is absent in the DADS–EDP record and earnings are positive, so the worker-year contributes neither to stock counts nor to any flow pair involving that year.

Differences are ≤ 0.03 pp/year for every period, occupation, and channel. The routine Excess_NE sign-flip is preserved: the pre-polarization value moves from +0.25 (imputed) to +0.22 (no imputation), the polarization value from −0.09 to −0.08, so the swing is −0.34 versus −0.30 — same sign, same order of magnitude. The bilateral routine-to-abstract flow (−0.52 to −0.54 across both configurations) and the manual non-employment entry channel (+0.25 to +0.24 during polarization) are likewise unchanged.

Because Section 5.1 highlights the education margin, we repeat the comparison stratified by college status (Tables 46–48).

Table 46: Education-stratified bilateral decomposition with CS2 imputation applied versus the missing-CS2 worker-years dropped — full panel 1984–2021 (pp/year).

Group	**Sample**	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	Imputed	manual	—	−0.13	−0.04	−0.18	+0.24	+0.06
		routine	+0.13	—	−0.53	−0.39	−0.00	−0.40
		abstract	+0.04	+0.53	—	+0.57	−0.24	+0.33
	No imp.	manual	—	−0.13	−0.04	−0.17	+0.23	+0.06
		routine	+0.13	—	−0.53	−0.40	+0.01	−0.39
		abstract	+0.04	+0.53	—	+0.57	−0.24	+0.33
Non-college	Imputed	manual	—	−0.13	−0.04	−0.17	+0.31	+0.14
		routine	+0.13	—	−0.45	−0.32	+0.00	−0.32
		abstract	+0.04	+0.45	—	+0.49	−0.31	+0.18
	No imp.	manual	—	−0.13	−0.04	−0.16	+0.30	+0.14
		routine	+0.13	—	−0.45	−0.33	+0.01	−0.32
		abstract	+0.04	+0.45	—	+0.49	−0.31	+0.18
College	Imputed	manual	—	−0.13	−0.10	−0.23	+0.19	−0.04
		routine	+0.13	—	−1.27	−1.14	+1.02	−0.12
		abstract	+0.10	+1.27	—	+1.37	−1.21	+0.16
	No imp.	manual	—	−0.13	−0.10	−0.22	+0.17	−0.06
		routine	+0.13	—	−1.30	−1.17	+1.04	−0.13
		abstract	+0.10	+1.30	—	+1.39	−1.21	+0.18

Notes: Values ×100 (pp/year). "Imputed" uses the CS2 imputation procedure of Section 2; "No imp." sets the occupational category to missing for every worker-year with raw CS2 absent and positive earnings.

Both education groups are likewise insensitive to imputation: ΔE_k, Net_JtJ, and Excess_NE entries differ by ≤ 0.04 pp/year in the polarization period and by ≤ 0.15 pp/year in the pre-polarization period (the larger pre-polarization variation reflecting the small back-projected education-observed sample). Taken together, the aggregate, non-college, and college results of Sections 4 and 5.1 are insensitive to whether the missing-CS2 worker-years are imputed or dropped.

Table 47: Education-stratified bilateral decomposition with CS2 imputation applied versus the missing-CS2 worker-years dropped — pre-polarization 1984–1993 (pp/year).

Group	**Sample**	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	Imputed	manual	—	−0.31	−0.04	−0.35	+0.20	−0.15
		routine	+0.31	—	−0.58	−0.27	+0.17	−0.10
		abstract	+0.04	+0.58	—	+0.62	−0.37	+0.26
	No imp.	manual	—	−0.31	−0.04	−0.35	+0.20	−0.15
		routine	+0.31	—	−0.58	−0.27	+0.16	−0.12
		abstract	+0.04	+0.58	—	+0.62	−0.35	+0.27
Non-college	Imputed	manual	—	−0.32	−0.04	−0.36	+0.23	−0.12
		routine	+0.32	—	−0.47	−0.14	+0.17	+0.02
		abstract	+0.04	+0.47	—	+0.50	−0.40	+0.10
	No imp.	manual	—	−0.32	−0.04	−0.36	+0.23	−0.12
		routine	+0.32	—	−0.46	−0.14	+0.16	+0.02
		abstract	+0.04	+0.46	—	+0.50	−0.39	+0.11
College	Imputed	manual	—	−0.10	−0.12	−0.22	+0.13	−0.09
		routine	+0.10	—	−2.17	−2.07	+1.62	−0.45
		abstract	+0.12	+2.17	—	+2.29	−1.75	+0.53
	No imp.	manual	—	−0.12	−0.11	−0.23	+0.11	−0.12
		routine	+0.12	—	−2.23	−2.11	+1.55	−0.57
		abstract	+0.11	+2.23	—	+2.34	−1.66	+0.68

Notes: Same sample definitions as Table 46.

Table 48: Education-stratified bilateral decomposition with CS2 imputation applied versus the missing-CS2 worker-years dropped — polarization 1994–2021 (pp/year).

Group	**Sample**	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	Imputed	manual	—	−0.08	−0.04	−0.12	+0.25	+0.13
		routine	+0.08	—	−0.51	−0.43	−0.06	−0.49
		abstract	+0.04	+0.51	—	+0.55	−0.20	+0.36
	No imp.	manual	—	−0.07	−0.04	−0.11	+0.24	+0.13
		routine	+0.07	—	−0.51	−0.44	−0.04	−0.48
		abstract	+0.04	+0.51	—	+0.55	−0.20	+0.35
Non-college	Imputed	manual	—	−0.07	−0.03	−0.11	+0.33	+0.22
		routine	+0.07	—	−0.45	−0.38	−0.05	−0.43
		abstract	+0.03	+0.45	—	+0.48	−0.28	+0.20
	No imp.	manual	—	−0.07	−0.03	−0.10	+0.32	+0.22
		routine	+0.07	—	−0.45	−0.38	−0.04	−0.42
		abstract	+0.03	+0.45	—	+0.49	−0.28	+0.20
College	Imputed	manual	—	−0.14	−0.10	−0.24	+0.21	−0.03
		routine	+0.14	—	−0.98	−0.84	+0.82	−0.02
		abstract	+0.10	+0.98	—	+1.08	−1.03	+0.05
	No imp.	manual	—	−0.13	−0.09	−0.22	+0.18	−0.04
		routine	+0.13	—	−1.00	−0.86	+0.88	+0.01
		abstract	+0.09	+1.00	—	+1.09	−1.06	+0.02

Notes: Same sample definitions as Table 46.

C.5 Robustness to Educational Data Missingness

Section 5.1 stratifies the sample by education using a worker-level construction: each worker takes the education category recorded at her first observation, with a diploma field — recoded to the same categories — filling in where the education field is missing. Two distinct sources of missingness affect this construction. First, education is entirely unrecorded in DADS before 2002; pre-2002 education is therefore reconstructed by carrying each worker's first observed value backward, which restricts the pre-2002 by-education sample to workers who survive in the panel until at least 2002. Second, even after 2002 a non-trivial share of employed workers has no observed education at any point: 4% in 2002, rising monotonically to 18% by 2021. Pooled across the panel, the workers with never-observed education account for the difference between the all-with-education black line and the full-sample grey line in Figure 8; that gap is small in all three occupation panels and in all years (at most 1.7 percentage points), confirming that aggregate within-employment occupational composition is not materially biased by the education-observation selection. The remaining question, which this section addresses, is whether the *by-education* results are sensitive to how the never-observed-education workers would be classified.

We construct a maximal-imputation bound under the assumption that every worker with never-observed education is non-college. This is the worst case for the non-college panel: the non-college pool expands by roughly 469k unique workers (around 33% over the 1.23M observed non-college pool), with their occupational distribution determining the bound. Symmetric imputation toward college is empirically less defensible: workers who never have education recorded over the panel are most plausibly drawn from lower-education cohorts, short-tenure workers, and panel boundary cases, all of which correlate with non-college status. We therefore report only the one-sided bound on non-college; the observed college line is interpreted as the closest available point estimate (any imputation toward college would push the line *downward*, but by an empirically negligible amount).

Within-employment shares. Figure 12 reports the bounded version of Figure 8, with the non-college imputation envelope shown as a shaded blue band (lower edge: observed non-college; upper edge: observed non-college + all never-observed-education workers imputed as non-college). The pre-2002 portion of the band reflects back-projection uncertainty (workers exiting before 2002 with never-observed education); the post-2002 portion reflects contemporaneous educ-observation gaps. The band has a U-shape in time: widest pre-2002 (at most 2.9 pp on Routine and 1.9 pp on Abstract in 1984) and post-2015 (climbing to 2.9 pp on Routine and 3.7 pp on Abstract in 2021), and essentially closed in 2002–2010 (< 1 pp). The qualitative patterns documented in Section 5.1 — a clear U-shape of polarization within non-college and a considerably flatter distribution within college — are robust to imputation throughout the panel.

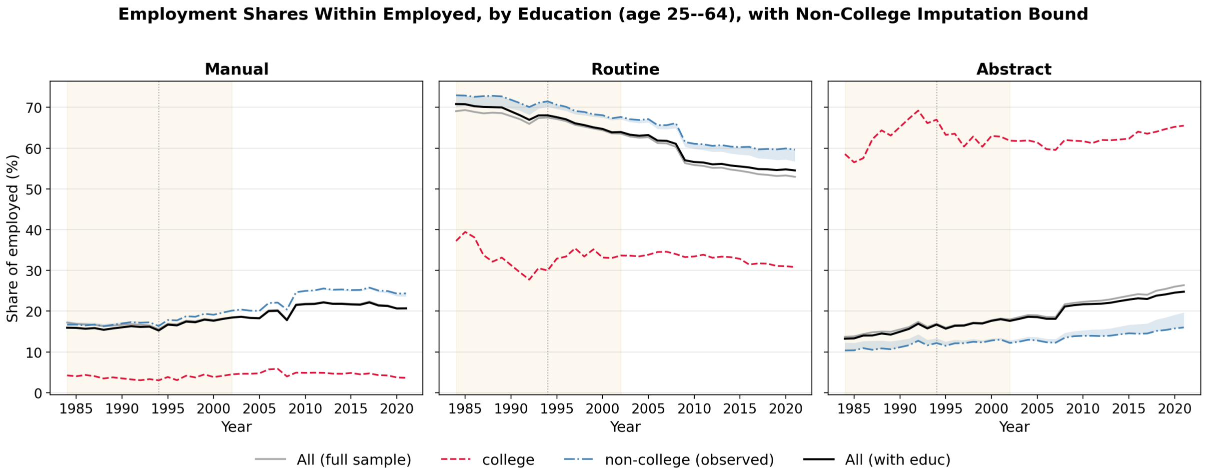

Figure 12: Within-employment occupational shares by education group, with the maximal-imputation bound on non-college shares shaded. Lower edge of the band: non-college as observed (Figure 8). Upper edge: non-college pool expanded to include all workers with never-observed education. Grey line: full age-25–64 sample. The pre-2002 portion of the band captures back-projection uncertainty (panel exits before 2002); the post-2002 portion captures observation gaps.

Bilateral flow decomposition. Table 49 reports the period-averaged bilateral decomposition for non-college workers under both the observed and the maximal-imputation samples. The pre-polarization and polarization sub-periods match the specification used in Figures 9 and 10 of the main text.

Table 49: Bilateral decomposition for non-college workers: observed (Section 5.1 baseline) vs. maximal-imputation bound (all never-observed-education workers imputed as non-college). Values are pp/year ×100. b_R, b_A, b_M: bilateral Net_JtJ contributions from Routine, Abstract, Manual. "—" denotes a self-flow excluded by construction.

		Observed non-college					Imputed non-college					
Period	**Occ**	b_R	b_A	b_M	ExNE	ΔE	b_R	b_A	b_M	ExNE	ΔE	$\Delta(\Delta E)$
Pre-pol 1984–93	Manual	−0.322	−0.037	—	+0.234	−0.125	−0.223	−0.027	—	+0.048	−0.203	−0.078
	Routine	—	−0.466	+0.322	+0.165	+0.021	—	−0.461	+0.223	+0.390	+0.153	+0.132
	Abstract	+0.466	—	+0.037	−0.399	+0.103	+0.461	—	+0.027	−0.438	+0.050	−0.054
Pol 1994–2021	Manual	−0.070	−0.035	—	+0.329	+0.224	−0.078	−0.042	—	+0.307	+0.187	−0.037
	Routine	—	−0.449	+0.070	−0.049	−0.428	—	−0.466	+0.078	−0.091	−0.479	−0.051
	Abstract	+0.449	—	+0.035	−0.280	+0.204	+0.466	—	+0.042	−0.215	+0.292	+0.088

Three patterns emerge. First, ΔE values shift by at most 0.13 pp/year under imputation, and in the polarization period by at most 0.09 pp/year. The non-college time-series patterns documented in Section 5.1 are well within these bounds. Second, the bilateral Net_JtJ flows (b_R, b_A, b_M) are remarkably stable across the two scenarios, with differences of at most 0.10 pp/year and concentrated on the manual-to-routine flow (b_M in the Routine row, b_R in the Manual row); the dominant routine-to-abstract drain (b_A in the Routine row) is essentially unchanged. Third, the routine Excess_NE channel exhibits the sign flip across the 1994 break under both specifications and the flip is in fact *stronger* under imputation: the observed-sample swing from +0.165 pre-pol to −0.049 pol becomes +0.390 to −0.091 under maximal imputation. The qualitative finding that the non-employment buffer feeding routine disappears around 1994 is therefore robust — and arguably understated, rather than overstated, by the observed-only specification.

C.6 The Non-Employment Channel and the Routine-Decline

Section 4.2 showed that the routine decline runs mainly through the non-employment channel. This appendix quantifies how much, and shows that the result holds across polarization subperiods and does not rest on the 2007–09 crisis years.

The routine decline accelerates after 1994: ΔE_r falls from −0.09 pp/year before the break to −0.52 over 1994–2021, an acceleration of −0.43. Table 50 attributes that acceleration to the two channels. In the full window the non-employment (Excess_NE) swing is −0.34 and the job-to-job swing only −0.09, so the Excess_NE swing accounts for 78% of the acceleration. The split is stable across windows: 79% in the pre-crisis 1994–2006 phase, and 83% once the 2007–09 pairs are dropped. The job-to-job upgrading drain is nearly constant across 1994 and in every window the channel's *swing* carries about four-fifths of the acceleration.

Table 50: Decomposition of the routine-decline acceleration by polarization window (pp/year). The routine swing is the sum of a job-to-job (Net_JtJ) swing and a non-employment (Excess_NE) swing. The last window drops the 2007–09 pairs.

Polarization window	ΔE_r	ΔE_r swing	Net_JtJ swing	Excess_NE swing	ExNE share
Full 1994–2021	−0.52	−0.43	−0.09	−0.34	78%
Pre-crisis 1994–2006	−0.44	−0.35	−0.07	−0.28	79%
Excl. 2007–09	−0.36	−0.26	−0.04	−0.22	83%

Following the analysis of Section 4, Table 51 splits each window's swing into an entry-deviation part (routine's entry deviation $\rho_r^{\text{ent}} - 1$ moving against its share-neutral benchmark), an exit-deviation part, and a benchmark part (the benchmark $E_r \sum_k F$ shrinking). The entry-deviation part dominates in every window (−0.16 to −0.31): routine's entry deviation widens from −0.03 to −0.05 or −0.06 — routine is increasingly passed over in non-employment hiring relative to its benchmark. The exit-deviation part is small and holds near −0.06 throughout so that routine's exit propensity relative to benchmark barely moves. The large exit-side adjustment that a glance at the channel might suggest is almost entirely a benchmark effect — the benchmark shrinking — not routine releasing workers faster.

Because the 2009 DADS Grand Format coincides with the Great Recession, the crisis years are a natural concern; the crisis-excluded window above shows the result does not

Table 51: Entry, exit and the routine Excess_NE swing.

Polarization window	entry $\rho_r - 1$	exit $\rho_r - 1$	Deviation, entry	Deviation, exit	Benchmark
Full 1994–2021	$-0.03 \to -0.06$	$-0.06 \to -0.06$	-0.31	$+0.01$	-0.04
Pre-crisis 1994–2006	$-0.03 \to -0.04$	$-0.06 \to -0.05$	-0.16	-0.02	-0.10
Excl. 2007–09	$-0.03 \to -0.05$	$-0.06 \to -0.06$	-0.21	$+0.06$	-0.07

rest on them. The Grand Format itself is taken up through the $(2008, 2010)$ bridge in Appendix C.3 and the imputation-drop exercise in Appendix C.4.