The Anatomy of Polarization

Evidence from Worker Flows

Fabio Cerina[*] Elisa Dienesch[†]
Alexander Monge-Naranjo[‡] Alessio Moro[§]

This version: 12 May 2026. For the latest version, click here.

Abstract

Using longitudinal French administrative data (1984–2021), we document that employment polarization after 1994 reflects major changes in labor-market entry rather than mass occupational downgrading or displacement of incumbents. Flows from routine to abstract occupations remain substantial throughout the period, and a large fraction of these upgrades is due to non-college workers. The decisive shift that generates polarization occurs at the entry margin: the net flow from non-employment into routine occupations reverses around 1994, while the net flow from non-employment into manual work increases. These patterns motivate life-cycle models of occupational choice that explicitly incorporate cohort heterogeneity and separate entry and re-entry margins.

Keywords: *Labor market polarization; Worker flows; Occupational mobility; Routine-biased technological change; Longitudinal data.*

JEL Codes: J62, J24, J21.

[*] University of Cagliari and CRENoS.
[†] University of Cagliari, CRENoS, and AMSE.
[‡] Federal Reserve Bank of Atlanta, Emory University, and CEPR.
[§] University of Cagliari, CEPR and IZA@LISER.

We thank participants in the Aix-en-Provence seminars and Cagliari workshops (Brucchi Luchino; “Inequalities and Labor Markets”) and 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 thanks 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 over 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). The literature has mostly relied on repeated cross-sectional evidence, which is silent about the underlying worker flows that drive polarization. However, the way workers move across occupational categories matters because the same decline in routine employment can either reflect a mass downgrading of incumbent routine workers into manual work and non-employment (as in Autor and Dorn, 2013), or be driven by new entrants avoiding routine jobs. Despite this, the role of worker flows in generating polarization has received little attention (Cortes et al., 2020). This paper fills this gap by exploiting exhaustive longitudinal French administrative data to quantify how much each margin (job-to-job flows, entry and exit) accounts for the observed employment polarization.

We characterize workers' reallocation flows using a longitudinal matched employer–employee panel for France that follows individual workers from 1984 to 2021. The panel covers all salaried employment (*tous salariés*), records each worker's annual occupation and non-employment spells, letting us trace careers across nearly four decades. 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. Using longitudinal data allows us to answer three basic questions regarding the emergence of polarization in France: whether the workers most affected are entrants or incumbents; whether career advancement across different occupation categories is accelerated or blocked; and whether measured employment polarization emerges through downgrading, through a failure of non-employment to replenish middle-skill jobs, or through enhanced upward mobility from the bottom to the top of the occupational ladder.

Aggregate employment shares show no clear trend before 1994: the routine share is broadly flat. After 1994 the pattern changes sharply: the routine share declines steadily while abstract and manual employment expand. We therefore split the sample into a pre-polarization phase (roughly 1984–1993) and a polarization phase (1994–2021). This break coincides with the sharp fall in information and communication technology (ICT) costs documented in the literature, which accelerated ICT adoption across sectors. We then identify which flows changed across these regimes. Our central finding is that the cross-sectional appearance of polarization is not driven by widespread downward displacement of routine incumbents; rather, it is driven by a collapse in the net 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.

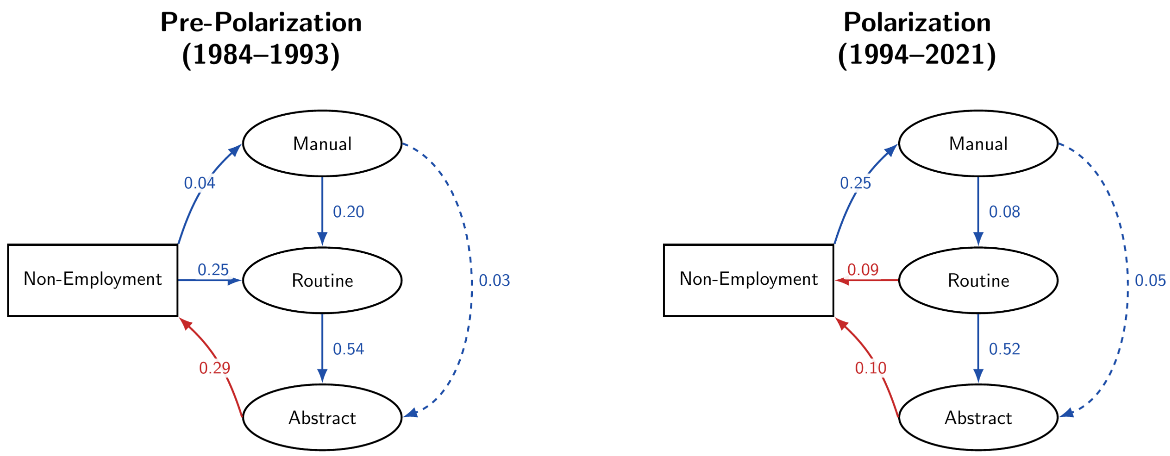

Figure 1: Net yearly worker flows in the pre-polarization (1984–1993) and polarization (1994–2021) regimes in percentage points. Blue arrows mark upgrading and entry-from-non-employment; red arrows downgrading or outflows to non-employment.

Within Routine itself, this entry-margin collapse is concentrated in the cognitive component — 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 therefore a routine-cognitive phenomenon specifically — the routine type most directly displaceable by ICT equipment.

A second key result from the changes in the flows is that the routine decline does not operate through downgrading into manual jobs or through displacement into non-employment. 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 gradual decline 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 simple displacement interpretation of routine-biased technical change; in the French microdata the displacement channel is absent.

A third finding is that employment polarization is a compositional effect across the education margin. 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 (+5.6 and +7.7 pp respectively).[1] About half of the 14-percentage-point rise in the abstract share over 1984–2021 reflects within-non-college upgrading; within-college upgrading contributes roughly one-sixth and the rising college share the remainder. The non-college contribution is overwhelmingly routine-to-abstract job-to-job upgrading. 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: 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, undergo monotone upgrading: their routine and manual shares both decline (−15 and −5 pp) while abstract gains over 20 pp. The decline of the college manual share is concentrated in the polarization era and is driven by stronger upgrading flows toward routine and weaker entry from non-employment relative to non-college workers.

Worker-level conditional analysis reinforces these patterns. Education is the dominant predictor of upgrading and downgrading: holding age, gender, and tenure fixed, college roughly triples the routine→abstract upgrading rate and substantially reduces downgrading from abstract. Occupation-specific tenure stabilizes workers, so polarization runs through entrants and short-tenured workers; 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

[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, and the gap is proportionally larger for non-college workers.

Loebbing, 2026) and, more directly, to work that emphasizes worker-level flows behind aggregate polarization. Closest to our approach, Cortes et al. (2020) use matched CPS data to decompose the U.S. routine employment decline and show 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. They pursue a flow decomposition using innovative methods, but the rotating CPS panel limits direct tracking of individual transitions over time. Our French DADS–EDP panel, which follows individual workers continuously across four decades, complements their evidence by recovering the job-to-job margin and showing that it is concentrated in routine-to-abstract upgrading rather than routine-to-manual displacement.

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 terminal state.[2] Bocquet (2024) also uses French DADS data to study occupational reallocation, modelling it as a network of feasible transitions with bridge occupations as bottlenecks; his question is why reallocation is slow, whereas ours is which margins drive polarization in the first place. We are not aware of previous works charachterizing employment polarization using job flows at such level of detail.

Methodologically, our contribution is to integrate entry into and exit from non-employment into the 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.

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.

[2] 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.

After sample restrictions the panel covers roughly 1.9 million workers over 1984–2021 (1990 is absent because of the census). DADS-EDP links annual DADS wage records to the Permanent Demographic Sample (EDP), a birth-date stratified sample of the French population.[3] Coverage includes both private and public salaried employment (*tous salariés*), giving DADS-EDP broader scope than DADS-Postes, which is restricted to the private sector.

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.[4] We assign a main occupation as the spell with the largest share 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

[3] 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.

[4] 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 year), with a gradient-boosted classifier as fallback. Details and a no-imputation robustness check are in Appendix D.4.

The classification is task-based rather than wage-ranked. It 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. 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 5).

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.

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 $k \in \{a, m, r, o\}$ to denote any of these states, and use e for the 'employment' state, i.e., the aggregate of workers in $k \in \{a, m, r\}$.[5]

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 independent cross-sectional snapshots of

[5]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 practising 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 D.3. Under the data-anchored bias correction implemented in that appendix, the qualitative findings of the paper are preserved and, if anything, *reinforced*.

the 25–64 window at each date.[6]

Let $N_e(t) \equiv N_m(t) + N_r(t) + N_a(t)$ denote total employment in year t. The *within-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 INSEE employment rate $e(t)$ for the 25–64 age group as an external denominator:[7]

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

These four shares sum to one: $S_m(t) + S_r(t) + S_a(t) + S_o(t) = 1$. The total working-age population is then $N(t) = N_e(t)/e(t)$, and the count 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 count of non-employed individuals, which is less well identified in the panel than the employment counts by occupation.

3.2 Worker Flows and Employment Shares

We now consider the worker flows across the four states. Let $F_{k,k'}(t)$ denote the number of workers in state k at $t-1$ and in state k' at t. The 4×4 flow matrix $[F_{k,k'}]$ 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 employment counts and worker flows are linked by the accounting identity: for all k and all t,

$$N_k(t) \;=\; N_k(t-1) + \sum_{k' \neq k} \Big[F_{k',k}(t) - F_{k,k'}(t)\Big]. \tag{4}$$

[6] So workers must be aged 25–64 at $t-1$ for the $t-1$ count and aged 25–64 at t for the t count. Boundary cohorts (aging-in at 24 → 25, aging-out at 64 → 65) are handled by the residual recovery of $F_{o,k}(t)$ in equation (5); the coding convention is in Appendix A.3.

[7] 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: $e(t) = \frac{25}{40}\, e_{25\text{–}49}(t) + \frac{15}{40}\, e_{50\text{–}64}(t)$.

The increment in the number of workers in occupation k equals the net inflow from all other states. Separating active reallocation flows from those that involve entry or exit through non-employment, we can *recover* the implied entry into occupation k from non-employment o as

$$F_{o,k}(t) \;=\; \Delta N_k(t) + F_{k,o}(t) - \text{Net\_JtJ}_k(t), \tag{5}$$

where $\Delta N_k(t) \equiv N_k(t) - N_k(t{-}1)$ is the observed change in employment in occupation k and

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

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

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.[8]

3.3 Decomposing Changes in Employment Shares

We now decompose the observed changes in within-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[9]

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

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.

From equation (4), the change in the count of state k equals the sum of net flows from all other states: $N_k(t) - N_k(t{-}1) = \sum_{k' \neq k}\Big[F_{k',k}(t) - F_{k,k'}(t)\Big]$. We separate these flows into a job-to-job component (*Net_JtJ*) and a non-employment component (*Net_NE*). For each employed occupation $k \in \{m, r, a\}$, the job-to-job component captures the net reallocation of incumbent workers into k,

$$\text{Net\_JtJ}_k(t) \;\equiv\; \sum_{k' \notin \{k,o\}} \Big[F_{k',k}(t) - F_{k,k'}(t)\Big],$$

[8] The alternative of *pooling* would give more weight to years with larger numbers of observations.

[9] An alternative decomposition is $\Delta E_k(t) = [\Delta N_k(t) - E_k(t{-}1)\cdot\Delta N_e(t)]/N_e(t)$, which uses the lagged share $E_k(t{-}1)$ as the benchmark.

and the non-employment component captures the net flow between k and non-employment (entries minus exits),

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

Substituting these definitions into the share-change formula yields

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

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

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

We also decompose the *Net_JtJ* component into bilateral net flows with each of the other employed occupations. For routine, in particular, we split the net JtJ flows into manual (downgrading) or into abstract occupations (upgrading).

4 Decomposing the Polarization Pattern

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 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.

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

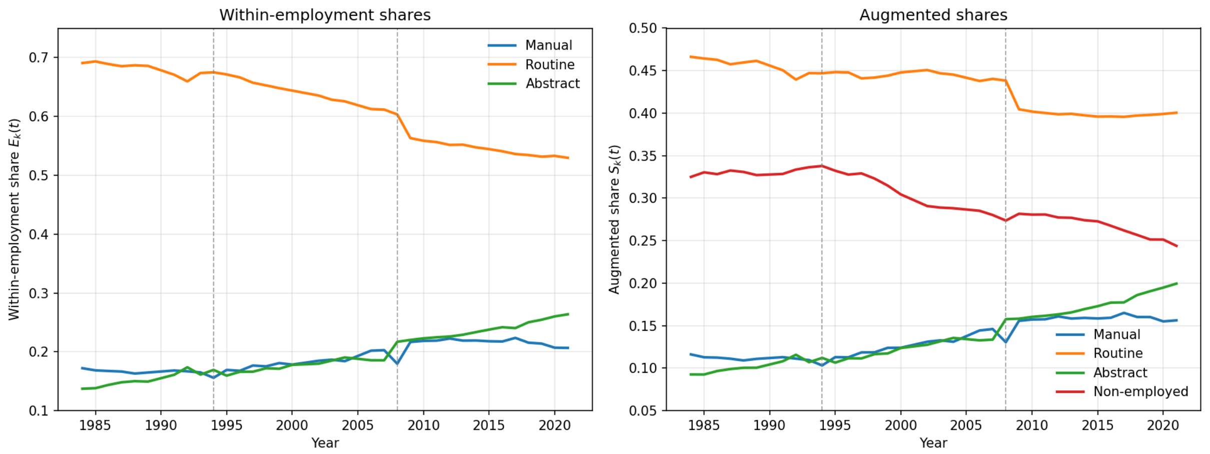

Figure 2: Occupational employment shares, workers aged 25–64. *Left:* within-employment shares $E_k(t)$. *Right:* augmented shares $S_k(t) = E_k(t) \cdot e(t)$ together with the non-employed share $S_o(t) = 1 - e(t)$, using the INSEE 25–64 employment rate. Dashed vertical lines mark 1994 and 2008.

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.[10] 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.

[10]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.

Table 1: Mean annual ΔE_k by sub-period (pp/year, simple-average, age 25–64)

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 $\Delta E_a = +0.40$. The post-1994 pattern fits the standard notion of employment polarization closely.

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 full-period picture can be summarized in three facts. 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. The manual rise ($\Delta E_m = +0.03$ pp/year) is driven by 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. 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 C.

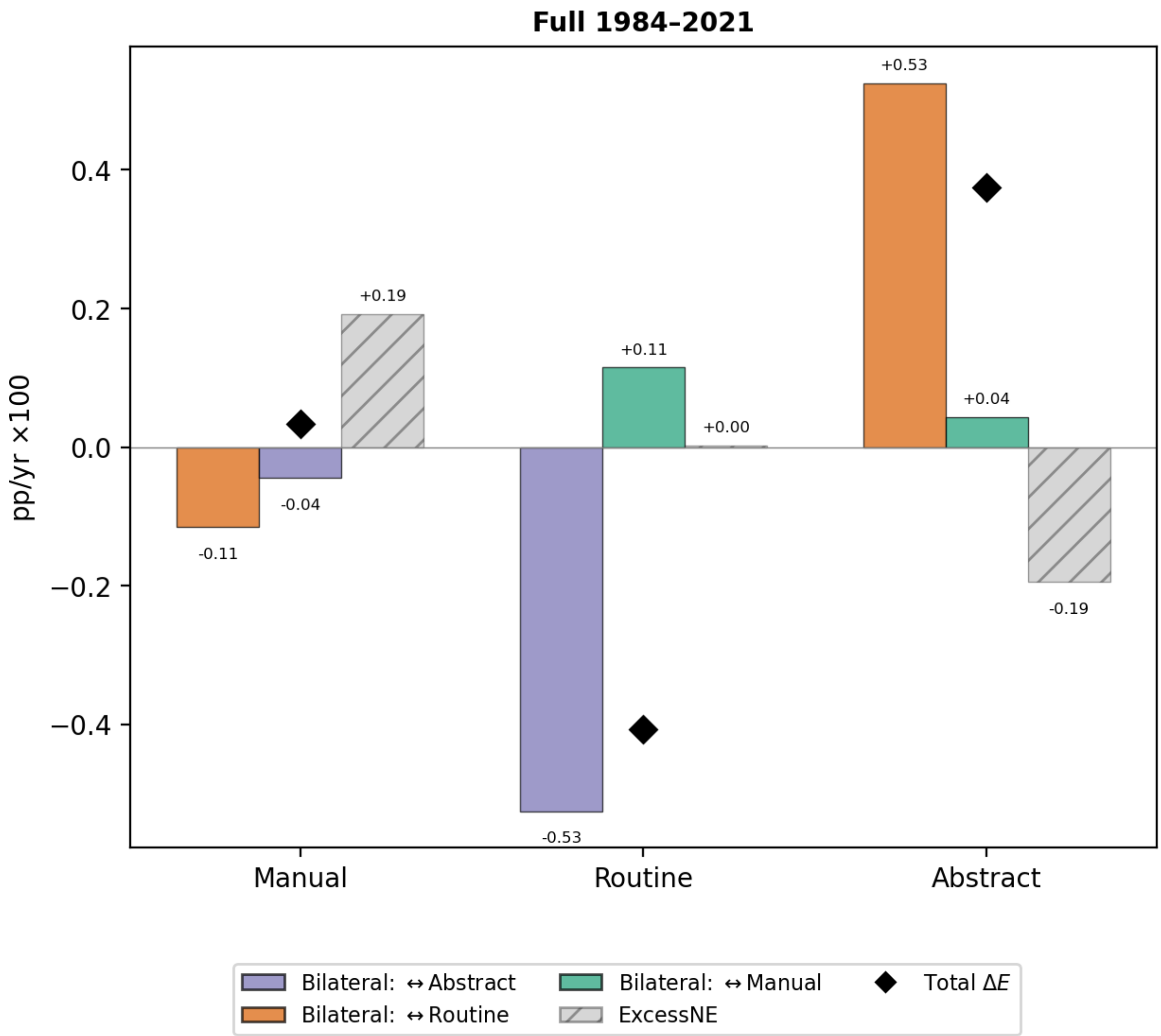

Figure 3: Mean annual bilateral Net_JtJ decomposition together with the Excess_NE channel and the resulting ΔE_k, full period 1984–2021. For each destination occupation k, the three solid bars show the bilateral net flows from the other two employed states; the hatched bar is Excess_NE; the diamond is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

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

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-j" reports the bilateral net inflow $(F_{j,k} - F_{k,j})/N_e(t-1)$; Net_JtJ is the sum of the three bilateral entries. Each panel sums to zero across occupations by construction. Bold marks the routine Excess_NE sign flip across the 1994 break. A finer sub-periodization (acute polarization 1994–2006, post-crisis 2009–2019, and the polarization regime excluding the crisis years) is reported in Appendix C.

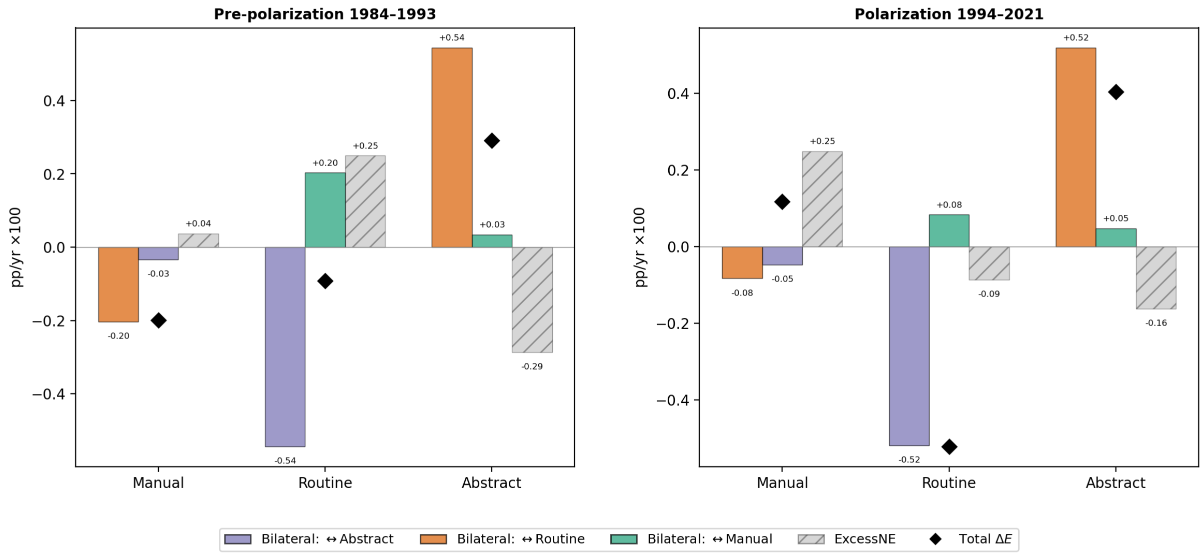

Figure 4: Mean annual bilateral decomposition: pre-polarization (1984–1993) vs. polarization (1994–2021).

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 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 Excess_NE swing is concentrated in the acute phase 1994–2006 ($\text{Excess\_NE}_r = -0.03$, already close to zero), deepens during the 2007–09 crisis (the crisis pairs pull the 1994–2021 mean down to -0.09), and partially recovers in the post-crisis phase ($\text{Excess\_NE}_r = +0.08$ over 2009–2019). A more robust statement than "sign flip" is that the non-employment buffer *disappears* around 1994 and remains either absent or mildly negative thereafter, with a transitory crisis episode on top. The fine sub-period decomposition is reported in Appendix C; 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 D.1).

Anatomy of the routine Excess_NE sign flip. 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. Table 3 decomposes the routine Excess_NE channel into an *entry residual* and an *exit residual*. Writing the Excess_NE channel as $\text{Excess\_NE}_r = R_r^{\text{ent}} + R_r^{\text{exit}}$, the entry residual is $R_r^{\text{ent}} = [F_{o,r} - E_r \sum_k F_{o,k}]/N_e$ (the $o \to r$ inflow net of its share-weighted benchmark), and the exit residual is $R_r^{\text{exit}} = [E_r \sum_k F_{k,o} - F_{r,o}]/N_e$ (the share-weighted exits to non-employment net of actual $r \to o$ exits). A negative entry residual means routine is under-represented among non-employment-to-employment entries; a positive exit residual means routine workers are *sheltered* from exits to non-employment (fewer $r \to o$ exits than the routine share of employment would imply).

Across 1994, both raw rates contract in similar proportion: the $o \to r$ entry rate falls from 10.61% to 6.81% of employment per year (a 36% reduction) and the $r \to o$ exit rate falls from 8.93% to 5.70% (a 36% reduction). The two movements push routine employment in opposite directions: a fall in $o \to r$ reduces the routine stock through reduced hiring, while a fall in $r \to o$ raises it through fewer departures. To assess which channel drives the 1994 break, consider a counterfactual in which the $o \to r$ rate had remained at its pre-polarization value (10.61%) while everything else — the $r \to o$ contraction, the mild

Table 3: Entry-exit decomposition of routine Excess_NE by sub-period. Values ×100 (pp/year).

Period	Entry side $o \to r$ rate	Entry side R_r^{ent}	Exit side $r \to o$ rate	Exit side R_r^{exit}	Excess_NE$_r$
Pre-pol 1984–1993	+10.61	−0.29	+8.93	+0.54	+0.25
Pol 1994–2021	+6.81	−0.44	+5.70	+0.35	−0.09
Δ across 1994	−3.80	−0.15	−3.23	−0.19	−0.34

Notes: $o \to r$ rate is the conditional transition rate of NE workers transitioning to routine in the next period; R_r^{ent} is the entry-side residual after netting out $E_r(t) \cdot \Delta N_e$; analogous decomposition on the exit side.

weakening of the Net_JtJ channel, and the other flows — moved as observed. Under that scenario, ΔE_r would be approximately +1.2 pp/year: the routine share would not decline at all. The collapse of the $o \to r$ entry rate is therefore responsible for the −0.52 pp/year routine decline during polarization; the $r \to o$ contraction partially offsets it. Routine workers do not lose their jobs at higher rates across the 1994 break — exits fall, not rise — so the sign flip is not a displacement story. The accounting locates it on the hiring margin: employers fill fewer routine positions from the non-employment pool during polarization than before.

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 C.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 much smaller role.

4.3 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.[11]

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 cancelling 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.

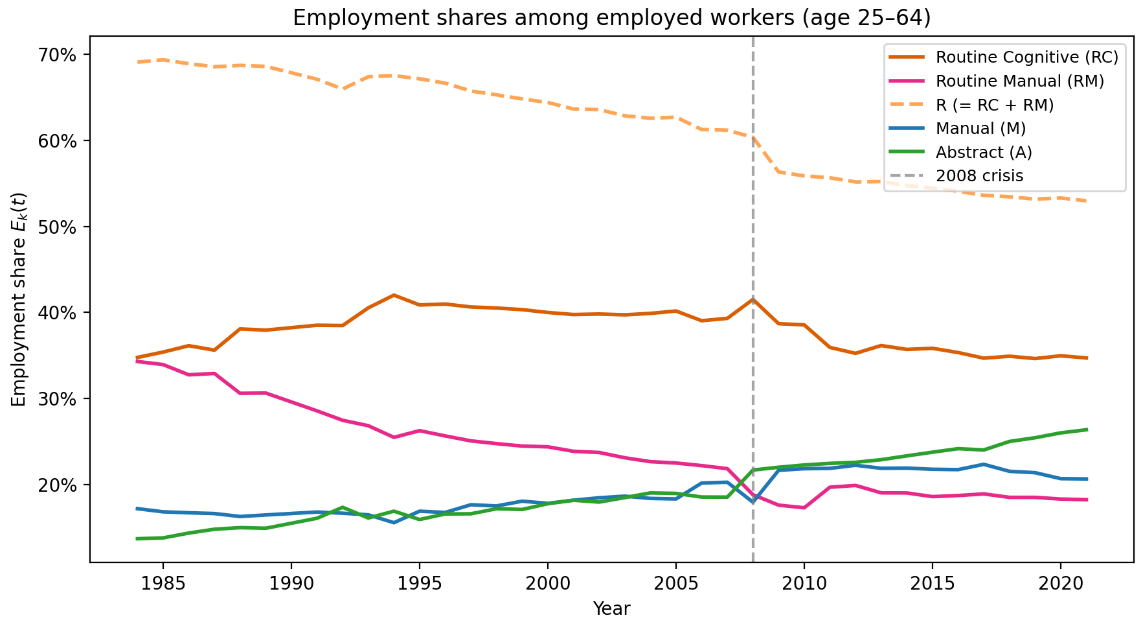

Figure 5: Within-employment shares, five-state system. The aggregate routine share (dashed) is the sum of rc (routine cognitive) and rm (routine manual).

4.3.1 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.[12]

[11]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).

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

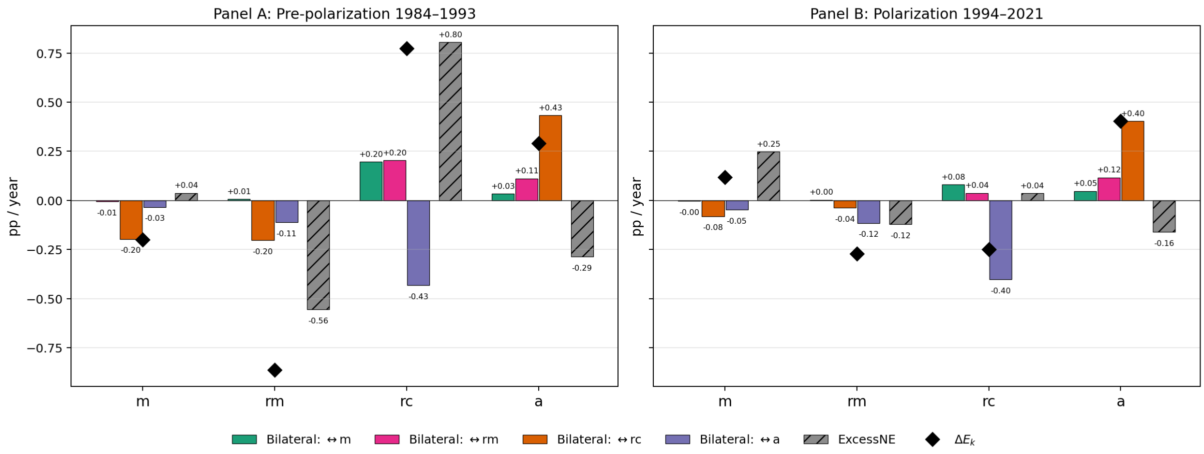

Figure 6: Mean annual bilateral decomposition of ΔE_k for the five-state system (m, rm, rc, a). Panel A: pre-polarization 1984–1993. Panel B: polarization 1994–2021. For each destination occupation, colored bars show the bilateral Net_JtJ flows from the other three employed states (green: $\leftrightarrow m$; 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 $m < rm < rc < a$.

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, i.e., rc *over*-represented in non-employment entries, to −0.25 during polarization), while rm's excess entry improves arithmetically 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 E, 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 work into manual routine production. The CS2 decomposition in Appendix E 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.

4.4 A candidate trigger: ICT prices

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. 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 around 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 4.3 are

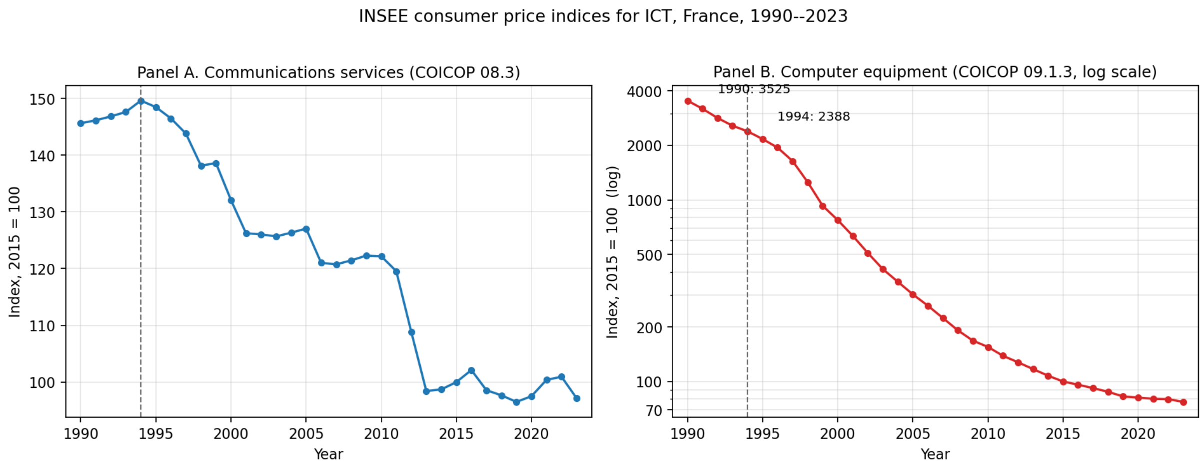

Figure 7: Consumer price indices for ICT, France, 1990–2023 (INSEE IPC, base 2015 = 100). Panel A: communications services (COICOP 08.3, idBank 001764308). Panel B: information processing equipment (COICOP 09.1.3, idBank 001764157), log scale.

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). The aggregate education stratification of these findings — whether rc workers are college-educated and whether the entry-margin collapse traces back to non-college workers — is taken up in the next Section 5.

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 behavior. 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 11.9% in 2002–05 to 17.1% in 2017–21. About 75% of unique workers in the panel have an education observation, covering about 85% 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. Among non-college workers the typical U-shape emerges: between 1984 and 2021 the routine share falls by 13.3 pp (from 72.4% to 59.1%) while both the manual share (+5.6 pp, from 19.5% to 25.1%) and the abstract share (+7.7 pp, from 8.1% to 15.8%) rise. College-educated workers, by contrast, undergo upgrading toward abstract occupations: the abstract share rises by 20.1 pp (from 45.4% to 65.5%), while both the routine share (−15.0 pp, from 45.7% to 30.7%) and the manual share (−5.2 pp, from 8.9% to 3.7%) decline.

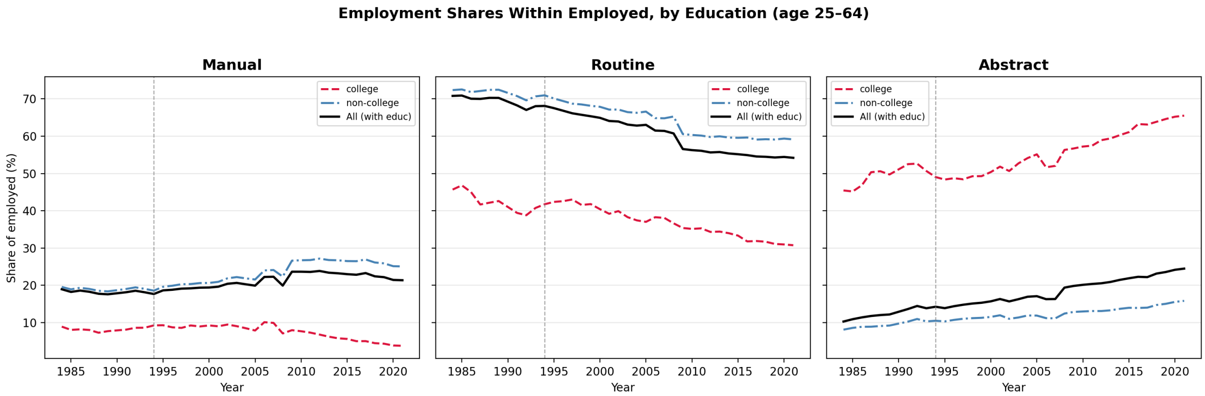

Figure 8: Within-employed occupational shares by education group. Each panel plots manual, routine, and abstract shares among employed college and non-college workers over 1984–2021.

An Oaxaca–Blinder decomposition of each aggregate share change between 1984 and 2021 quantifies the relative weight of within-group reallocation versus the rising college share (Appendix C.5.1; Table 15 reports the full breakdown). Importantly, about half of the +14-pp aggregate rise in the abstract share reflects within-non-college upgrading, while within-college upgrading contributes roughly one-sixth, and the rising college share the remainder. The aggregate routine decline (-16.6 pp) is overwhelmingly a within-non-college phenomenon (-11.7 pp, 70%), with smaller contributions from within-college routine decline (10%) and the rising college share (19%). The aggregate manual share rises modestly (+2.4 pp), but this masks a sizable within-non-college rise (+4.9 pp) that is partly offset by a within-college decline and by the rising college share itself.

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 C.5.

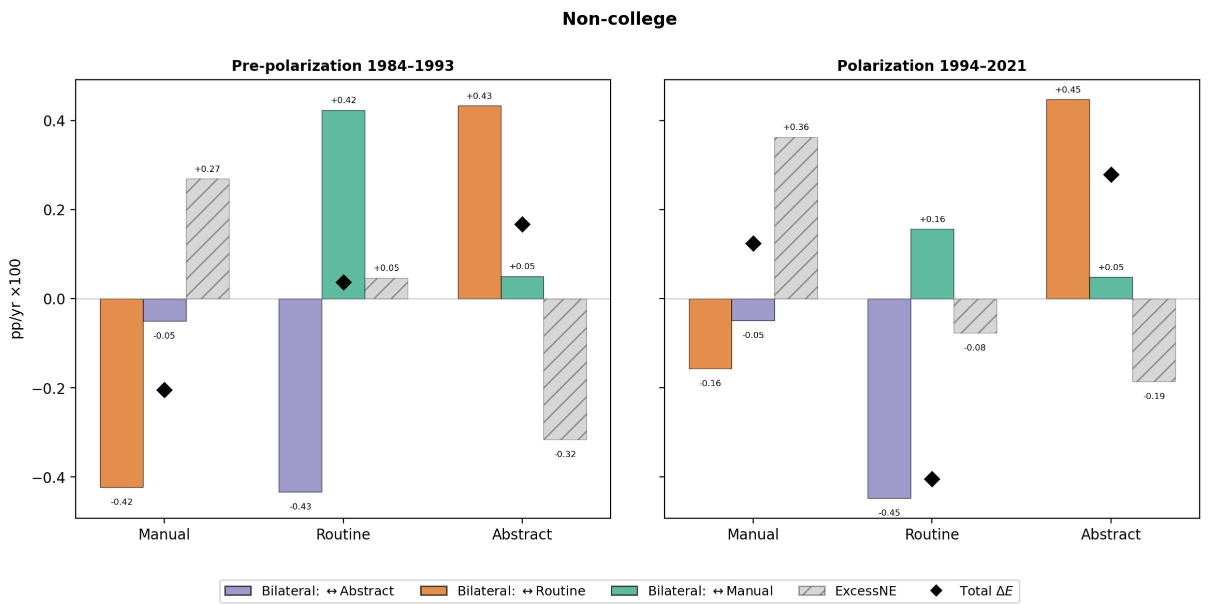

Figure 9: Mean annual decomposition for non-college workers, pre-polarization (left, 1984–1993) and polarization (right, 1994–2021, excluding DADS expansion pairs 2001–02 and 2005–06). Bars show bilateral Net_JtJ flows (net from m, r, a), the Excess_NE channel, and the resulting ΔE_k (diamond).

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 +0.43 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 +2.18 pp/year pre-polarization and +1.08 pp/year during polarization. The college rate is roughly five times the non-college rate pre-1994 and about 2.4 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 decomposition

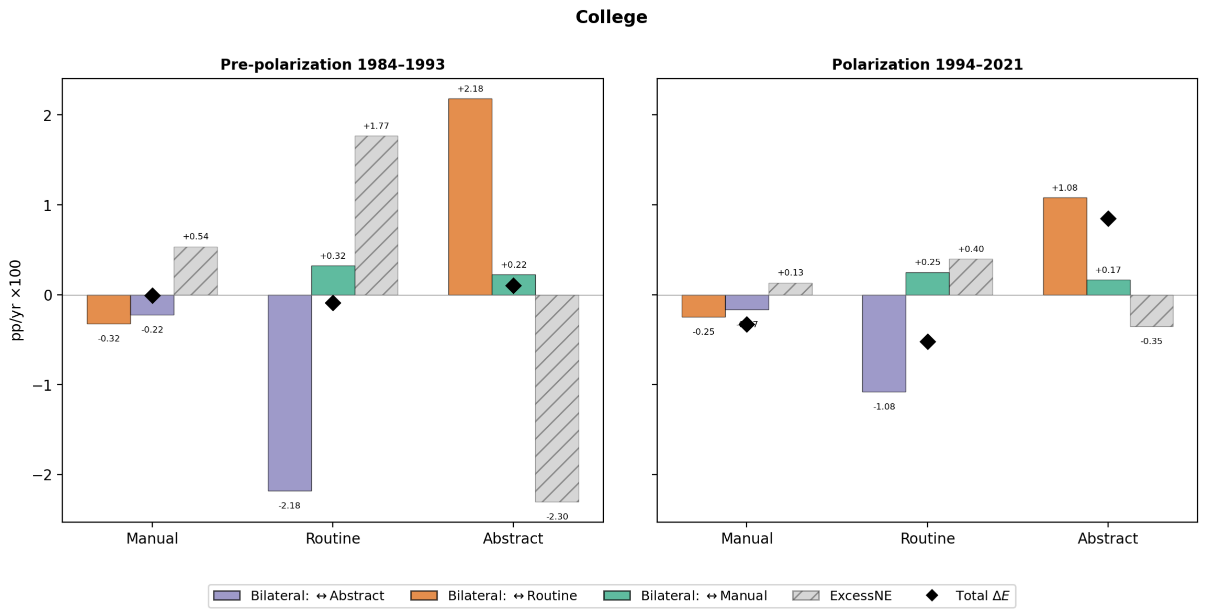

Figure 10: Mean annual decomposition for college workers.

above.

The cross-period change in ΔE_r inside the non-college sample, from +0.04 pp/year pre-polarization to −0.40 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 (+0.42 to +0.16 pp/year, a −0.27 swing) and the routine non-employment entry channel flips sign (+0.05 to −0.08 pp/year). The two together account for the full swing in non-college ΔE_r. Inside the college sample the routine decline accelerates from $\Delta E_r = -0.09$ to −0.52 pp/year, driven primarily by a falling non-employment-to-routine entry rate.

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 −0.20 pre-polarization to +0.12 pp/year during polarization, a positive swing of +0.32. The bilateral $r \to m$ flow inside the non-college sample is, however, negative in both periods (net $m \to r$ is +0.42 pp/year pre-polarization and +0.16 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 +0.36 pp/year, almost three times the corresponding college value of +0.13 (Appendix C.5).

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

College status may confound other worker attributes such as age, gender, and tenure. To isolate the partial contribution of each characteristic we estimate linear probability models (LPMs) for one-year transitions from each origin state $j \in \{m, r, a, o\}$ to each destination $k \neq j$. Define binary transition indicators $y_{it}^{j\to k} = 1$ if worker i in state j at t is in state k at $t{+}1$. Our baseline specification is

$$y_{it}^{j\to k} = \alpha + \beta_1\,\mathrm{age}_c + \beta_2\,\mathrm{female} + \delta\,\mathrm{tenure} + \mathrm{Year\ FE} + \varepsilon_{it}, \tag{6}$$

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. The reason mirrors the discussion in Sections 2 and 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 EDP sample, given the frequent jumps in panel size from the 2002 and 2006 expansions. For these o-origin rows, tenure in equation (6) is replaced by the duration of the current observed non-employment spell ending at t (in years), with the standard re-employment hazard interpretation; the years-in-current-PCS interpretation applies only to the employed-origin rows. We report three nested models. Specification S1 (baseline, no education) uses all 35 matched year-pairs and estimates equation (6). Specification S2 adds a binary college indicator (Licence or higher) and restricts the sample to post-2002 observations where education is observed.[13] Specification S3 adds a female×college interaction to S2.[14]

[13]After worker-level forward/backward fill of education across the panel, post-2002 coverage exceeds 95% for both Net_JtJ flows and the o-origin re-entrants sample. Coefficients on age, female, and tenure are stable between the full-sample S1 and the post-2002 subsample S2, so restricting to post-2002 is not a first-order concern.

[14]All qualitative conclusions are robust to replacing the LPM with a probit estimated on the same specification.

Table 4 reports college, female, and tenure coefficients from specification S2 for the twelve one-year transitions, organised into four groups: within-employment upgrading, within-employment downgrading, exits to non-employment, and (re-)entries from non-employment. Full regression tables for the upgrading and re-entry transitions, including S1 and S3, appear in Appendix F (Tables 29–34).

Table 4: LPM partial correlations by transition type (specification S2, post-2002): all twelve transitions

Transition	$\bar{y}$	**College**	**Female**	**Tenure**
Within-employment upgrading				
$m \to r$	0.103	$+0.058^{***}$	-0.028^{***}	-0.008^{***}
$r \to a$	0.032	$+0.062^{***}$	-0.020^{***}	-0.003^{***}
$m \to a$	0.012	$+0.050^{***}$	-0.010^{***}	-0.001^{***}
Within-employment downgrading				
$r \to m$	0.038	-0.013^{***}	-0.004^{***}	-0.003^{***}
$a \to r$	0.061	-0.050^{***}	$+0.015^{***}$	-0.006^{***}
$a \to m$	0.008	-0.008^{***}	$+0.000^{***}$	-0.001^{***}
Exits to non-employment				
$a \to o$	0.070	-0.013^{***}	$+0.001$	-0.005^{***}
$r \to o$	0.090	-0.007^{***}	-0.004^{***}	-0.008^{***}
$m \to o$	0.128	$+0.001$	$+0.012^{***}$	-0.012^{***}
(Re-)entries from non-employment (re-entrants only)				
$o \to m$	0.334	-0.242^{***}	$+0.094^{***}$	$+0.011^{***}$
$o \to r$	0.511	-0.145^{***}	$+0.005^{***}$	-0.008^{***}
$o \to a$	0.155	$+0.388^{***}$	-0.099^{***}	-0.003^{***}

Notes: Each row reports coefficients from a separate LPM regression of the indicated transition on a college dummy, a female dummy, age (linear), tenure, and year fixed effects, with HC3-robust standard errors. For employed origins ($j \in \{m, r, a\}$) tenure is years in the current two-digit PCS code; for o-origin rows the sample is restricted to re-entrants and tenure is the duration of the current observed non-employment spell ending at t. $\bar{y}$ is the unconditional probability of the transition within the origin-state sample. Sample is the post-2002 window with observed education (worker-level forward/backward fill applied to recover education from any observed year). $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Full regressions (including the full-sample S1 without college and the S3 extension with a female×college interaction) for the upgrading and re-entry transitions are in Appendix F.

First, *education is the strongest predictor of upgrading.* For the $r \to a$ transition — the main margin behind aggregate polarization — the college coefficient is +6.2 pp on an unconditional baseline of 3.2%. Holding age, gender, tenure, and year fixed, college roughly triples the annual probability of an $r \to a$ move. An even larger college premium appears for $o \to a$ re-entries (+38.8 pp on a 15.5% baseline), the largest effect in the table; evaluated at the non-college baseline, college roughly quintuples the abstract re-entry rate. A similar direction holds for $m \to a$ (+5.0 pp on a 1.2% baseline). Symmetrically, college workers in abstract are 5.0 pp less likely to downgrade to routine (baseline 6.1%) and 1.3 pp less likely to exit to non-employment (baseline 7.0%); the corresponding effects on routine exits (-0.7 pp on a 9.0% baseline) and manual exits (+0.1 pp on a 12.8% baseline) are smaller and not always of the expected sign. Education is therefore associated with both faster upgrading into abstract and slower downgrading from abstract, while its effects on exits to non-employment are modest.

Second, *tenure is a strong stabilizer.* A one-year increase in occupation-specific tenure lowers the $r \to a$ probability by 0.30 pp, the $m \to r$ probability by 0.80 pp, and the $a \to r$ probability by 0.60 pp. As workers accumulate occupation-specific human capital they are less likely to move up or down the ladder, which reinforces the view that polarization operates mainly through entrants and relatively young workers rather than through reshuffling of long-tenured incumbents. Age coefficients are omitted from Table 4: once tenure is included they are typically less than 15% of the tenure effect in absolute magnitude. The visible exceptions are $a \to o$ and $r \to o$, where age and tenure carry opposite signs — older workers in abstract or routine jobs are more likely to exit to non-employment, consistent with end-of-career retirement, while shorter tenure independently predicts higher exit risk.

Third, *gender matters on upgrading margins.* The female coefficient for $r \to a$ is negative and large (-2.0 pp on a 3.2% baseline), so women are substantially less likely than men to upgrade from routine to abstract at a given age, tenure, and year. The same pattern appears for $m \to a$ (-1.0 pp) and $o \to a$ re-entries (-9.9 pp; the largest gender gap in the table) and for $a \to r$ (+1.5 pp; positive, so women are more likely to downgrade). On the (re-)entry side into routine and manual from non-employment, the female coefficient is positive ($o \to r$: +0.5 pp; $o \to m$: +9.4 pp), so women's disadvantage is concentrated on upgrading to abstract rather than on (re-)entries into routine and manual. Adding a female×college interaction (S3, Appendix F) attenuates the college premium for women on the three upgrading margins ($r \to a$, $m \to a$, $o \to a$: female×college coefficients of -4.1, -2.3, and -8.8 pp). On the two routine destinations ($m \to r$ and $o \to r$), by contrast, the female-specific college effect is positive (+4.2 and +18.0 pp), so college-educated women are disproportionately more likely than college-educated men to move into or re-enter routine occupations. This is consistent with rising female participation feeding into routine rather than directly into abstract.

Taken together, the evidence complements the aggregate picture in Sections 4: the college premium on $r \to a$ survives conditioning on age, gender, tenure, and year and is quantitatively large; the manual margin of polarization is largely driven by younger and less-tenured workers; and gender operates asymmetrically, with women disadvantaged on upgrading to abstract but not on extensive-margin entries into routine.

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 flow from non-employment into routine: entry rates into routine roughly halve and non-employment ceases to be a net feeder of the middle-skill routine occupations. We also find that the rise in manual employment is driven mainly by entrants from non-employment disproportionately flowing into manual jobs. 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 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.

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.

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 elevator: The spatial dimension of job polarization in france. 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.

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.

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 5 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×, and routine manual roughly 1.3–1.5×. 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). This supports treating the *rc*/*rm* split as a genuine structural divide.

A.2 Employment Spells and NE Gaps

An *employment spell* is a maximal sequence of consecutive years with `occ_0` $\in \{1, 2, 3\}$. An *NE interruption gap* is a maximal sequence of consecutive years with `occ_0` $= 0$, 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 (≥ 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 count $N_k(t)$, 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 (4), 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 (5) absorbs aging-in flows

Table 5: CS2 codes ranked by mean daily wage relative to manual average

			1984		**1994**		**2009**		**2019**	
CS2	**Description**	**Group**	Rk	$w/\bar{w}_m$	Rk	$w/\bar{w}_m$	Rk	$w/\bar{w}_m$	Rk	$w/\bar{w}_m$
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 aged 25–64 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).

automatically: they enter $\Delta N_k(t)$ through the t snapshot but contribute to neither JtJ 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.

An alternative method directly cross-tabulates all $F_{k,k'}$ from the matched panel using a nuanced pre-entry coding. The two approaches produce identical JtJ flows, exits, and recovered entries for every consecutive-year pair; they differ only in the $o \to o$ cell, which does not enter the employment-share decomposition. Details are in Appendix D.2.

Workers first entering the 25–64 window at age 25 number 468,000 (28.5% 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 (+5 to +7 pp manual premium, no trend), 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) and are manual-heavy (33.4%) and routine-light (37.6%).

Boundary flows by construction affect only the non-employment channel: JtJ flows require both origin and destination to be in employed states ($k \neq o$), 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 D.1).

A.4 Exit Patterns

We classify each employed→NE transition as *definitive* (the worker never returns to employment in the remaining panel window) or *temporary* (the worker re-enters at some later year). The two types have sharply different age profiles. Definitive exits show median ages of 47 ($m \to o$), 53 ($r \to o$), and 54 ($a \to o$), with 33–49% of definitive exiters aged ≥ 55. Temporary exits show median ages of 36–38 across all occupations. The 8–16 year gap in median age confirms that the panel separates retirement-like permanent exits from transitory career interruptions. The abstract group has the highest definitive-exit share (49%) and the oldest profile, consistent with longer careers.

The relative weight of definitive vs. temporary exits is not constant over time. Figure 11 plots the share of employed→NE transitions that are definitive, by origin occupation. The series are stable around 25–35% before the 2008 crisis, drift upward through the 2010s, and rise sharply at the end of the panel. The terminal jump is a mechanical artefact of the right-censoring rule: as the panel horizon shrinks, fewer post-exit re-entries can be observed, so exits in the last few years of the panel are increasingly classified as definitive. The ranking across occupations is preserved throughout: the abstract group has the highest definitive share in every year, consistent with its older exit-age profile.

Figure 11: Share of employed→NE transitions classified as definitive, by origin occupation and year. Definitive exits are those with no re-entry observed in the remaining panel window. The terminal upturn reflects right-censoring of the panel.

B 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 k at $t+1$ conditional on being in state j at t; rows sum to one. Each cell of an unconditional matrix gives the probability of the (j, k) transition in the population; the full matrix sums to one.

B.1 Conditional Matrices

Table 6: Conditional matrix — full period 1984–2021.

From j \ To k	o	m	r	a
o	0.759	0.074	0.141	0.026
m	0.138	0.717	0.134	0.011
r	0.099	0.039	0.829	0.033
a	0.091	0.009	0.089	0.811

Table 7: Conditional matrix — pre-polarization 1984–1993.

From j \ To k	o	m	r	a
o	0.712	0.070	0.190	0.027
m	0.167	0.633	0.188	0.012
r	0.126	0.045	0.793	0.037
a	0.130	0.011	0.142	0.717

Table 8: Conditional matrix — polarization 1994–2021.

From $j \setminus$ To k	o	m	r	a
o	0.776	0.075	0.123	0.026
m	0.127	0.747	0.115	0.011
r	0.089	0.037	0.843	0.031
a	0.077	0.008	0.071	0.845

The key regime change is visible in the $o \to r$ cell: from 0.190 in the pre-polarization phase to 0.123 during polarization, a 35% decline. $o \to m$ is stable ($0.070 \to 0.075$), 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 (0.037 pre-pol, 0.031 pol).

B.2 Unconditional (Joint) Matrices

Table 9: Unconditional matrix — full period 1984–2021

From $j \setminus$ To k	o	m	r	a	Row sum
o	0.277	0.027	0.052	0.010	0.365
m	0.018	0.095	0.017	0.001	0.132
r	0.039	0.016	0.323	0.013	0.391
a	0.010	0.001	0.009	0.092	0.112
Col sum	0.344	0.138	0.402	0.116	1.000

Table 10: Unconditional matrix — pre-polarization 1984–1993

From $j \setminus$ To k	o	m	r	a	Row sum
o	0.263	0.026	0.070	0.010	0.369
m	0.019	0.072	0.021	0.001	0.113
r	0.054	0.019	0.342	0.016	0.431
a	0.011	0.001	0.012	0.062	0.086
Col sum	0.347	0.118	0.446	0.089	1.000

The unconditional matrices reveal the population weight of each flow. The $o \to r$ cell drops from 0.070 (pre-pol) to 0.045 (pol), the largest off-diagonal decline in the matrix.

Table 11: Unconditional matrix — polarization 1994–2021

From j \ To k	o	m	r	a	Row sum
o	0.282	0.027	0.045	0.009	0.364
m	0.018	0.103	0.016	0.002	0.138
r	0.034	0.014	0.317	0.012	0.376
a	0.009	0.001	0.008	0.103	0.122
Col sum	0.343	0.146	0.386	0.126	1.000

The r row sum (which reflects routine's population share) declines from 0.431 to 0.376 across the two regimes, but the composition of flows out of routine shifts markedly: exits to non-employment fall (from 0.054 to 0.034) while the diagonal rises, consistent with higher within-state persistence among the shrinking routine population.

C 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.

C.1 Fine Sub-period Decomposition

Table 12 splits the polarization regime reported in Table 2 of the main text into three non-overlapping windows: an acute 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 labour market rather than a polarization-era phenomenon. Second, the acute 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 -0.03) in their cleanest form; this is the window in which the routine decline is largest ($\Delta E_r = -0.44$). 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.

Table 12: Detailed bilateral decomposition by sub-period (pp/year)

Period		Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
Acute polarization 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) and the DADS expansion pairs $(2001, 2002)$ and $(2005, 2006)$.

C.2 Five-state Decomposition with RC/RM Split

Table 13 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 4.3.

Table 13: Five-state mean annual decomposition (pp/year), age 25–64.

	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 $\in \{62, 63, 65, 67\}$).

C.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 14.

Table 14: 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 stand out. 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 12 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 documented in Section 4 for direct JtJ flows is preserved, not reversed.

C.4 Breakdown of the Routine Non-Employment Channel

Section 4.2 reports the headline result that the routine non-employment channel switches sign across regimes (positive in 1984–93, mildly negative or zero in the polarization period). Table 12 above shows that this swing is concentrated in 1994–2006 and partially recovers post-crisis. To unpack the mechanism behind these aggregates, Figure 13 decomposes the

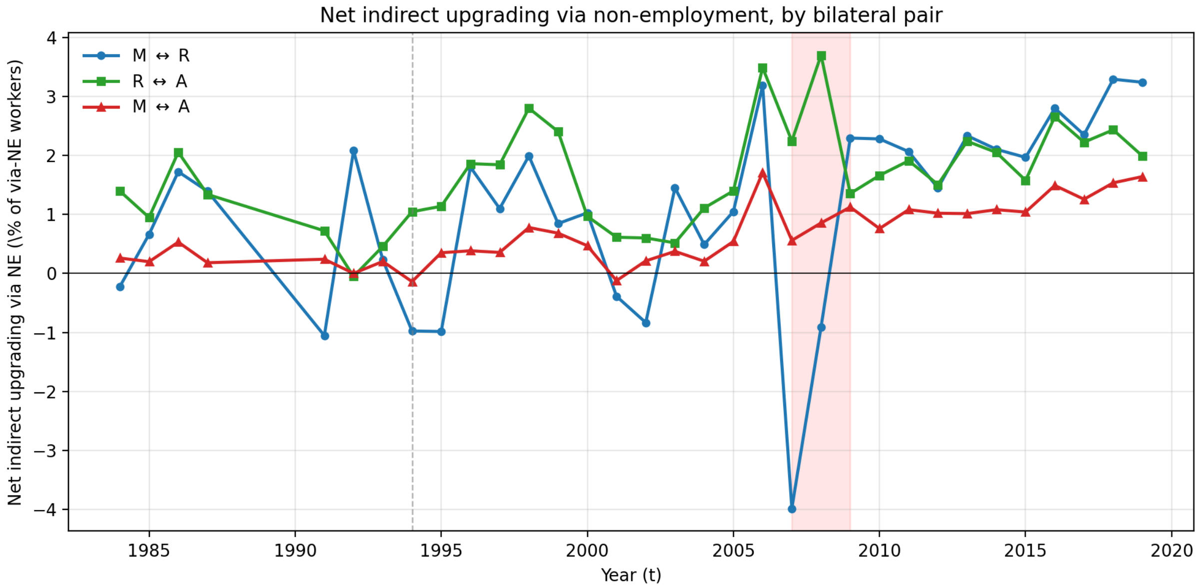

Figure 12: Net indirect upgrading via non-employment, by bilateral pair and year. For each pair (k, ℓ), the line reports the unconditional share of via-o workers following $k \to o \to \ell$ minus the share following $\ell \to o \to k$. Positive values indicate net upgrading.

routine non-employment channel into its entry and exit components.

Four facts emerge. *First*, the destination composition of $o \to$ employed entries (panel a) shifts sharply against routine over the panel: the routine share of $o \to$ employed inflows falls from about 65% in the mid-1980s to roughly 50% by the late 2010s, while the abstract share roughly doubles from ∼10% to ∼25% and the manual share is broadly flat. *Second*, the origin composition of employed $\to o$ exits (panel b) tells a similar story but more muted: the routine share of exits drifts down from ∼65% to ∼50%, while the abstract share rises to about 30%. The two compositional shifts together imply that routine becomes a smaller share of *both* sides of the non-employment turnover, but its share of inflows falls faster than its share of outflows. *Third*, panels (c) and (d) decompose the routine non-employment channel into its entry and exit residuals, R_r^{ent} and R_r^{exit} defined in Table 3 of the main text. The sub-period averages reconcile with that table: pre-polarization $R_r^{\text{ent}} = -0.29$, $R_r^{\text{exit}} = +0.54$, sum $= +0.25 \approx \text{ExNE}_r^{\text{pre}}$ (Table 2); polarization $R_r^{\text{ent}} = -0.45$, $R_r^{\text{exit}} = +0.36$, sum $= -0.09 \approx \text{ExNE}_r^{\text{pol}}$. Panel (c) shows the year-by-year time profile of the same two residuals; panel (d) plots the sub-period means with a diamond marker for the algebraic sum. The contest between a sharply falling entry residual (R_r^{ent} moves from -0.29 to -0.44 across 1994) and a smaller exit-residual reduction (R_r^{exit} moves from $+0.54$ to $+0.35$) is what drives the routine sign-flip. The sign reversal of the routine non-employment channel is therefore an *entry-margin* phenomenon: it is driven by the collapse of $o \to r$ recruitment, not by an acceleration of $r \to o$ exits.

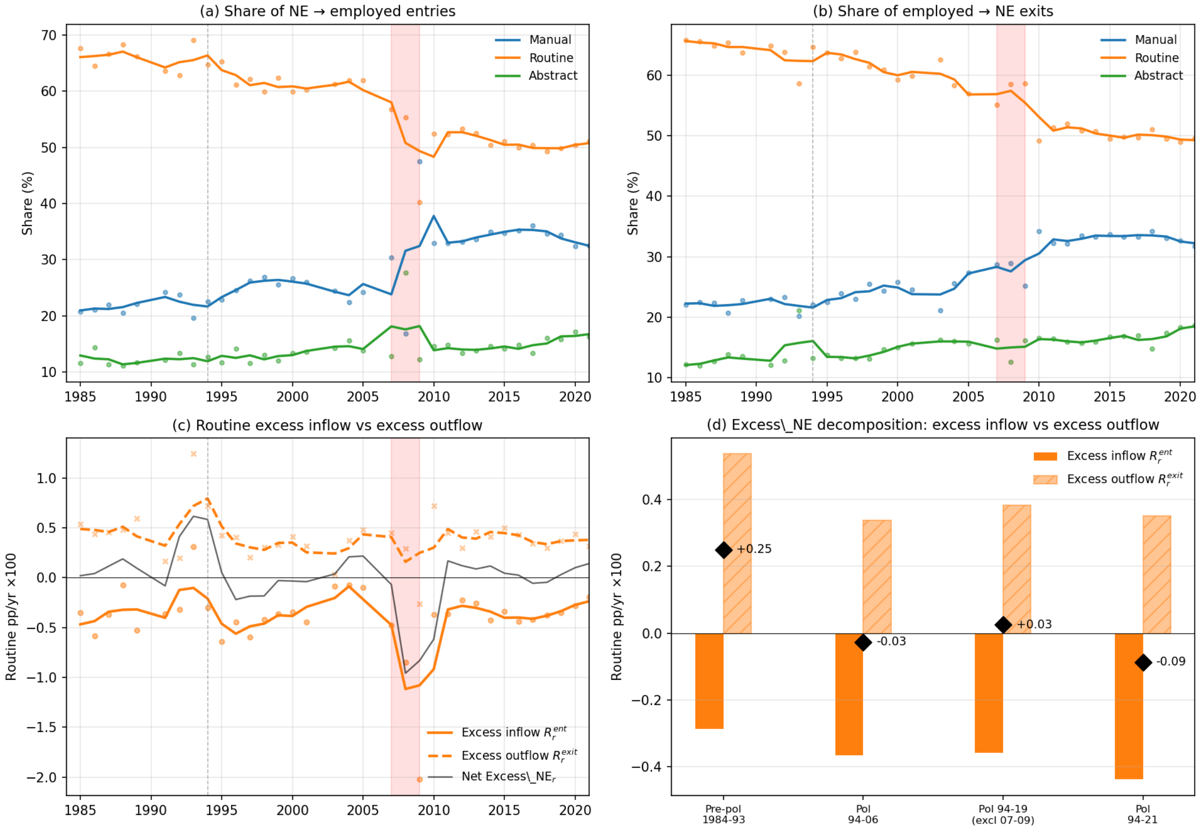

Figure 13: Breakdown of the non-employment channel. *Top row:* composition shares of o-mediated flows, by year. **Panel (a)**: share of $o \to$ employed entries $F_{o,k}$ going to each destination occupation $k \in \{m, r, a\}$; **Panel (b)**: share of employed $\to o$ exits $F_{k,o}$ coming from each origin occupation k. *Bottom row:* the routine non-employment channel decomposed into entry and exit residuals R_r^{ent} and R_r^{exit} defined in Table 3. **Panel (c)**: time series of the two residuals and of the resulting net $\text{Excess\_NE}_r = R_r^{\text{ent}} + R_r^{\text{exit}}$. **Panel (d)**: subperiod means of the same two components, with the diamond marker giving the net Excess_NE_r in each window. The shaded band flags the 2007–09 crisis years. Routine units throughout, in pp/year ($\times 100$).

C.5 Additional evidence on the education margin

This appendix collects supplementary evidence underlying the education-stratified results of Section 5.1. Subsection C.5.1 formalises the Oaxaca-Blinder decomposition of within-employed share changes by education group. Subsection C.5.2 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.

C.5.1 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 15 reports the decomposition for each occupational state over the full panel and the pre-polarisation and polarisation sub-periods.

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 manual share rises only modestly in aggregate but the within-non-college term is much larger (+4.9 pp out of an aggregate +2.4 pp), partly offset by a within-college decline and by the rising college share itself.

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 41–57% (midpoint 48%).

Table 15: 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	+2.4	+4.9	−0.6	−1.9
	Pre-pol 1984–1993	−0.8	−0.5	−0.0	−0.3
	Pol 1994–2021	+3.7	+5.6	−0.7	−1.2
Routine	Full 1984–2021	−16.6	−11.7	−1.7	−3.2
	Pre-pol 1984–1993	−2.8	−1.6	−0.4	−0.8
	Pol 1994–2021	−13.9	−10.2	−1.5	−2.2
Abstract	Full 1984–2021	+14.2	+6.8	+2.3	+5.0
	Pre-pol 1984–1993	+3.6	+2.1	+0.4	+1.1
	Pol 1994–2021	+10.2	+4.6	+2.2	+3.4

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 C.5.1). Sample restricted to employed workers aged 25–64 with education observed; college share among employed-with-education rises from 5.9% in 1984 to 17.4% in 2021.

C.5.2 Bilateral flows by education and period

Tables 16–18 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-j 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.

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 (+0.43 pp/year pre-polarization vs +0.45 post-1994), while inside the college sample it is large in both periods (+2.18 vs +1.08). The non-college routine decline during polarization therefore reflects the weakening of the $m \to r$ buffer flow (+0.42 to +0.16) and the sign flip of the non-employment entry channel into routine (+0.05 to −0.08), not a slowdown in $r \to a$ upgrading. Second, the manual share rise during polarization inside the non-college sample (ΔE_m = +0.12 pp/year, vs −0.20 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 +0.42 and +0.16 respectively). The non-college manual rise is instead absorbed by the manual non-employment entry channel (Excess_NE$_m$ = +0.36 pp/year during polarization).

Table 16: 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	—	−0.23	−0.06	−0.29	+0.28	−0.04
	routine	+0.23	—	−0.53	−0.31	−0.06	−0.39
	abstract	+0.06	+0.53	—	+0.60	−0.14	+0.43
Non-college	manual	—	−0.22	−0.05	−0.27	+0.34	+0.04
	routine	+0.22	—	−0.44	−0.22	−0.05	−0.30
	abstract	+0.05	+0.44	—	+0.49	−0.22	+0.25
College	manual	—	−0.27	−0.18	−0.45	+0.23	−0.25
	routine	+0.27	—	−1.35	−1.08	+0.73	−0.42
	abstract	+0.18	+1.35	—	+1.53	−0.82	+0.67

Notes: Values ×100 (pp/year), age 25–64, sample restricted to workers with education observed (worker-level forward/backward fill; rising college share among employed from 5.9% in 1984 to 17.4% in 2021). Net-j entries report the bilateral net JtJ flow into k from origin j. Expansion pairs $(2001, 02)$ and $(2005, 06)$ excluded.

Table 17: 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	—	−0.42	−0.06	−0.48	+0.25	−0.23
	routine	+0.42	—	−0.56	−0.14	+0.05	−0.08
	abstract	+0.06	+0.56	—	+0.62	−0.30	+0.31
Non-college	manual	—	−0.42	−0.05	−0.47	+0.27	−0.20
	routine	+0.42	—	−0.43	−0.01	+0.05	+0.04
	abstract	+0.05	+0.43	—	+0.48	−0.32	+0.17
College	manual	—	−0.32	−0.22	−0.55	+0.54	−0.01
	routine	+0.32	—	−2.18	−1.86	+1.77	−0.09
	abstract	+0.22	+2.18	—	+2.41	−2.30	+0.10

Notes: Values ×100 (pp/year), age 25–64. Sample as in Table 16.

Table 18: 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	—	−0.17	−0.06	−0.23	+0.29	+0.02
	routine	+0.17	—	−0.53	−0.36	−0.09	−0.49
	abstract	+0.06	+0.53	—	+0.59	−0.09	+0.47
Non-college	manual	—	−0.16	−0.05	−0.21	+0.36	+0.12
	routine	+0.16	—	−0.45	−0.29	−0.08	−0.40
	abstract	+0.05	+0.45	—	+0.50	−0.19	+0.28
College	manual	—	−0.25	−0.17	−0.42	+0.13	−0.33
	routine	+0.25	—	−1.08	−0.83	+0.40	−0.52
	abstract	+0.17	+1.08	—	+1.25	−0.35	+0.85

Notes: Values ×100 (pp/year), age 25–64. Sample as in Table 16. Expansion pairs $(2001, 02)$ and $(2005, 06)$ excluded.

D Robustness

D.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 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.

D.2 Alternative Identification of Non-Employment Entries

Section 3.2 introduces two alternative ways of identifying $o \to$ occupation entries: (i) recovering them as a residual from the accounting identity ($F_{o,k} = \Delta N_k - \text{Net\_JtJ}_k + F_{k,o}$), and (ii) directly cross-tabulating $F_{j,k}$ from the matched panel using a nuanced pre-entry coding. The two approaches yield identical JtJ flows, exits to non-employment, and recovered entries for every consecutive-year pair and every occupation. They differ only in the $o \to o$ cell $F_{o,o}$: approach (ii) counts it from the panel cross-tabulation, while approach (i) infers it from the INSEE employment rate. Since $F_{o,o}$ does not enter the employment-share decomposition, both approaches yield the same decomposition results.

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

The state o in our framework pools genuine non-employment, self-employment, informal work, and salaried positions outside DADS-EDP coverage. This raises two distinct (but interacting) concerns. First, the $o \to k$ and $k \to o$ flows include self-employment churn whose level, time-path, and cross-occupation distribution are non-trivial, even in the absence of any specific 2009 event. Second, the 2009 DADS-grand-format expansion (the auto-entrepreneur regime introduced 1 January 2009 plus the integration of CESU declarations) produces a stock jump in CS2=56 (manual category) in 2009. We address each in turn and conclude with the interaction between them.

D.3.1 Controlling for self-employment

Splitting $o = o_1 \cup o_2$ into genuine non-employment (o_1) and self-employment plus other DADS-uncovered states (o_2), and writing $\nu_k \equiv F_{o_2,k} - F_{k,o_2}$ and $\nu \equiv \sum_k \nu_k$, the Excess_NE channel decomposes as

$$\text{Excess\_NE}_k^{\text{obs}}(t) = \text{Excess\_NE}_k^{\text{NE-only}}(t) + \Xi_k(t), \quad \Xi_k(t) \equiv \nu_k(t) - E_k(t)\nu(t).$$

We discipline the inputs to Ξ_k in three sources:

1. **Aggregate** $\nu(t)$ from INSEE *Estimations d'emploi*, 1970–2018[15]: pre-pol 1984–93 average $\nu^{\text{pre}} = +81$ k/yr, pol-acute 1994–2008 average $\nu^{\text{pol}} = +11$ k/yr (the self-employment count declines ~ 80 k/yr pre-pol and ~ 11 k/yr pol-acute, then rebounds post-2009 with the auto-entrepreneur regime).
2. **Cross-occupation distribution** $\sigma_k(t)$ from INSEE Focus n°324 (1982–2022)[16] combined with the paper's PCS-1 to $\{m, r, a\}$ mapping. The series is monotonic and continuous across 1994: σ_r moves from 0.092 (1991) to 0.102 (1993) to 0.102

[15] https://www.insee.fr/fr/statistiques/4470794, Figure 2.
[16] https://www.insee.fr/fr/statistiques/8173452.

(1994) to 0.106 (1995). $\Delta\sigma_r$ across 1994 is +0.0002 — a near-zero discontinuity. The estimates are corroborated by URSSAF administrative data on *travailleurs indépendants* (2004–2024), which give σ_r in the 0.09–0.13 range.

3. **Annual transition rate** τ from salarié or chômeur to non-salarié, from Bauer–Garbinti–Georges-Kot (*INSEE Références* 2018)[17], LFS men 20–64: $\tau^{\text{pre}} = 1.4\%$ (1985–94), $\tau^{\text{pol}} = 0.8\%$ (1995–2004). This is the unconditional sample mean of the transition indicator.

JtJ flows. JtJ flow *counts* $F_{j,k}$ for $j, k \in \{m, r, a\}$ exclude self-employed workers (o_2), since self-employment is observed at neither side of the JtJ transition. The pooling does not affect counts.

JtJ *rates* $F_{j,k}/N_e^{\text{obs}}$ have two distortions. The denominator under-states true employment by the factor $1 - s(t)$ where $s(t)$ is the self-employment share; from the World Bank/ILO indicator `SL.EMP.SELF.ZS`,[18] s shifts smoothly from 0.152 (1991) to 0.114 (1995), giving a $\sim 2\%$ within-flow bias on rate levels — uniform across all bilateral rates, so cross-flow comparisons are unaffected. The misclassification distortion (a worker transitioning $j \to o_2 \to k$ enters as $j \to o$ rather than $j \to k$) is bounded by $F_{j\to k\,\text{via slf}}/N_e \le E_j \tau \sigma_k^{|j}$ where $\sigma_k^{|j}$ is the conditional probability of true occupation k given salaried-j becomes self-employed.

We discipline $\sigma_k^{|j}$ using the paper's own within-DADS conditional matrix (Tables 9 and 10), renormalised within $\{m, r, a\}$. Workers transitioning to self-employment plausibly behave *at least as occupation-stickily* as workers transitioning between occupations within DADS (a salaried plumber becomes a self-employed plumber, not a self-employed lawyer). The diagonal is high ($\sigma_r^{|r} = 0.91$ pre-pol, 0.93 pol), so off-diagonal biases are small. The bias-corrected routine row of the main text Table 2 shifts by at most 0.01 pp/yr:

	Observed			**Bias-corrected (within-DADS anchor)**		
	pre	pol	Δ	pre	pol	Δ
Net-m	+0.20	+0.08	−0.12	+0.19	+0.08	−0.11
Net-a	−0.54	−0.52	+0.02	−0.53	−0.52	+0.01
Net JtJ$_r$	−0.34	−0.44	−0.10	−0.34	−0.44	−0.10

For comparison, the absolute worst-case proportional allocation $\sigma_k^{|j} = \sigma_k$ (which implies a 69% probability that a routine clerk becoming self-employed becomes an artisan, agriculteur, or shopkeeper — empirically implausible) gives a bilateral correction of order 0.5 pp/yr. Under any data-anchored allocation consistent with the within-DADS conditional, the bilateral routine row is robust.

[17] https://www.insee.fr/fr/statistiques/3641409.

[18] https://api.worldbank.org/v2/country/FR/indicator/SL.EMP.SELF.ZS?format=json&per_page=80.

Excess_NE channel. Under proportional allocation $\nu_k = \sigma_k \nu$, the bias is

$$\Xi_r(t) = (\sigma_r - E_r)\nu/N_e.$$

Plugging in: $\Xi_r^{\text{pre}} = -0.21$ pp/yr, $\Xi_r^{\text{pol}} = -0.02$ pp/yr, $\boxed{\Delta\Xi_r = +0.19 \text{ pp/yr}}$. The sign is unambiguously *positive*: if any, the observed swing -0.34 pp/yr *understates* the true (NE-only) magnitude.

Under bias correction, the implied true swing is $-0.34 - 0.19 = -0.53$ pp/yr, so **larger** than the unbridged value. Under the absolute maximally-alarming but unrealistic proportional allocation, $|\Delta\Xi_r| \leq 0.23$ pp/yr, leaving the corrected swing no smaller than -0.11 pp/yr — still a sign-flip, still economically large.

The smooth time profile of $\sigma_r(t)$ across 1994 ($\Delta\sigma_r = +0.0002$) and the monotonic trajectory of $\nu(t)$ (smaller magnitude pol than pre) jointly rule out a 1994-discontinuity in Ξ_r that could mimic the observed sign-flip. For the routine sign-flip to be an artefact, Ξ_r would simultaneously need to swing by at least -0.34 pp/yr across 1994, be discontinuous at 1994 (in spite of monotonic σ_r, ν, τ trajectories), operate through a routine-specific channel (in spite of σ_r being the smallest of the three), and produce no parallel discontinuity in any French self-employment statistic at 1994 (none documented). We are not aware of any combination of mechanisms producing this.

Table 3 routine anatomy. The bias on the $o \to r$ rate (entry side) is bounded by $\sigma_r \cdot |\nu|/N_e \leq 0.13$ pp pre-pol, 0.10 pp pol-acute, with $\Delta \leq 0.04$ pp across 1994 — versus the observed shift -3.8 pp ($\sim 1\%$). The bias on the conditional $r \to o$ rate is bounded by $\tau \leq 1.4$ pp pre-pol, 0.8 pp pol-acute, with $\Delta \leq 0.6$ pp — versus the observed shift -3.2 pp ($\sim 19\%$). Table 3 is insensitive to the potential self-employment bias.

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

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.

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 20): Net JtJ$_r$

Table 19: 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

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 20: Bilateral net JtJ flows by sub-period (pp/year). Each cell reports $(F_{j,k} - F_{k,j})/N_e(t-1) \times 100$. Net JtJ_k is the row sum.

	Bilateral net inflow into k from			
k \ 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

Interaction between the two concerns. The two concerns are not fully independent. The 2009 auto-entrepreneur regime moves a substantial fraction of o_2 workers into DADS-salaried, mechanically reducing the size of o_2 relative to o post-2009 and therefore reducing its contamination on the post-2009 part of the polarization period. Combined with the (2008, 2010) bridge (which absorbs the level shift), the bridged metrics represent a sample in which both contaminations are attenuated. This interaction *strengthens* the joint robustness conclusion.

We conclude that the 1994 routine Excess_NE sign-flip is robust to: (i) general self-employment contamination (under the data-anchored within-DADS conditional, the bilateral row of Table 2 is essentially unchanged and the swing is *strengthened* to −0.53 pp/yr; under the maximally-alarming bound the swing is at most attenuated to −0.11 pp/yr but still negative); (ii) the 2009 DADS-grand-format event (the (2008, 2010) bridge attenuates the swing only marginally, from −0.34 to −0.30 pp/yr); and (iii) their joint structure (the post-2009 reduction of o_2 attenuates the self-employment bias on the bridged sample).

D.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 21 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

Table 21: 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.

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 stock counts 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 22 reports the main bilateral decomposition (Table 2 in the main text) under both samples.

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 23–25 reproduce the by-education decompositions of Tables 16–18 on both samples.

The non-college panel is essentially unchanged across the two samples: the bilateral $r \to a$ flow is −0.45 pp/year under both, the routine Excess_NE keeps a small negative sign in the polarization period (−0.08 vs −0.04), and the manual Excess_NE stays large and positive (+0.36 vs +0.32). The two regimes' qualitative contrast — steady $r \to a$ upgrading, weakening $m \to r$ buffer, sign change in the routine non-employment entry channel — is preserved.

The college sub-panel, on the other hand, shrinks visibly when the missing-CS2 worker-years are dropped: the polarization-era ΔE_a^c moves from +0.85 to +0.02, and ΔE_r^c from −0.52 to +0.01. This is exactly what should be expected given the composition of the

Table 22: Mean annual bilateral decomposition with CS2 imputation applied versus dropping the missing-CS2 worker-years (pp/year, age 25–64).

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. Bold marks the routine Excess_NE sign-flip across the 1994 break. Differences between the two samples are ≤ 0.03 pp/year for every cell.

Table 23: Education-stratified bilateral decomposition with CS2 imputation applied versus the missing-CS2 worker-years dropped — full panel 1984–2021 (pp/year, age 25–64).

Group	**Sample**	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	Imputed	manual	—	−0.23	−0.06	−0.29	+0.28	−0.04
		routine	+0.23	—	−0.53	−0.31	−0.06	−0.39
		abstract	+0.06	+0.53	—	+0.60	−0.14	+0.43
	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	−0.23	+0.33
Non-college	Imputed	manual	—	−0.22	−0.05	−0.27	+0.34	+0.04
		routine	+0.22	—	−0.44	−0.22	−0.05	−0.30
		abstract	+0.05	+0.44	—	+0.49	−0.22	+0.25
	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	−0.30	+0.18
College	Imputed	manual	—	−0.27	−0.18	−0.45	+0.23	−0.25
		routine	+0.27	—	−1.35	−1.08	+0.73	−0.42
		abstract	+0.18	+1.35	—	+1.53	−0.82	+0.67
	No imp.	manual	—	−0.13	−0.10	−0.22	+0.17	−0.06
		routine	+0.13	—	−1.30	−1.17	+1.10	−0.13
		abstract	+0.10	+1.30	—	+1.39	−1.17	+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. The smaller no-imputation ΔE_k magnitudes inside the College sub-panel reflect the predominance of FPE (*fonction publique d'État*) workers among the affected observations: dropping them removes a large share of the within-college flow signal, while the aggregate and non-college panels are essentially unchanged.

Table 24: Education-stratified bilateral decomposition with CS2 imputation applied versus the missing-CS2 worker-years dropped — pre-polarization 1984–1993 (pp/year, age 25–64).

Group	**Sample**	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	Imputed	manual	—	−0.42	−0.06	−0.48	+0.25	−0.23
		routine	+0.42	—	−0.56	−0.14	+0.05	−0.08
		abstract	+0.06	+0.56	—	+0.62	−0.30	+0.31
	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	—	−0.42	−0.05	−0.47	+0.27	−0.20
		routine	+0.42	—	−0.43	−0.01	+0.05	+0.04
		abstract	+0.05	+0.43	—	+0.48	−0.32	+0.17
	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	—	−0.32	−0.22	−0.55	+0.54	−0.01
		routine	+0.32	—	−2.18	−1.86	+1.77	−0.09
		abstract	+0.22	+2.18	—	+2.41	−2.30	+0.10
	No imp.	manual	—	−0.12	−0.11	−0.23	+0.11	−0.12
		routine	+0.12	—	−2.23	−2.11	+1.75	−0.57
		abstract	+0.11	+2.23	—	+2.34	−1.62	+0.68

Notes: Same sample definitions as Table 23. The shaky pre-polarization education-stratified numbers reflect both the small pre-2002 education-observed sample (workers must survive to 2002 to have education back-projected) and, in the no-imputation sample, the additional loss of FPE-concentrated worker-years.

Table 25: Education-stratified bilateral decomposition with CS2 imputation applied versus the missing-CS2 worker-years dropped — polarization 1994–2021 (pp/year, age 25–64).

Group	**Sample**	k	Net-m	Net-r	Net-a	Net_JtJ	Excess_NE	ΔE
All	Imputed	manual	—	−0.17	−0.06	−0.23	+0.29	+0.02
		routine	+0.17	—	−0.53	−0.36	−0.09	−0.49
		abstract	+0.06	+0.53	—	+0.59	−0.09	+0.47
	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	−0.19	+0.35
Non-college	Imputed	manual	—	−0.16	−0.05	−0.21	+0.36	+0.12
		routine	+0.16	—	−0.45	−0.29	−0.08	−0.40
		abstract	+0.05	+0.45	—	+0.50	−0.19	+0.28
	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	−0.27	+0.20
College	Imputed	manual	—	−0.25	−0.17	−0.42	+0.13	−0.33
		routine	+0.25	—	−1.08	−0.83	+0.40	−0.52
		abstract	+0.17	+1.08	—	+1.25	−0.35	+0.85
	No imp.	manual	—	−0.13	−0.09	−0.22	+0.19	−0.04
		routine	+0.13	—	−1.00	−0.86	+0.89	+0.01
		abstract	+0.09	+1.00	—	+1.09	−1.03	+0.02

Notes: Same sample definitions as Table 23. The bilateral $r \to a$ flow inside the non-college sample is essentially identical across the two samples (−0.45 under both), and the routine Excess_NE keeps a small negative sign (−0.08 vs −0.04). The college panel shrinks visibly when the missing-CS2 worker-years are dropped, because FPE-concentrated worker-years carry a large share of the within-college flow signal during polarization.

dropped observations. The worker-years in question are concentrated in FPE occupations — cadres (CS2 33–34), professeurs and instituteurs (CS2 42, 34), employés civils de la fonction publique (CS2 52, 53) — which are predominantly college-educated. The no-imputation sample is therefore not representative of the French labor market: it selectively excludes a large share of college-educated public-sector workers, who carry much of the within-college flow signal during polarization. The imputed sample reinstates this signal and is the appropriate baseline for analysis.

Taken together, the aggregate and non-college results of Sections 4 and 5.1 are insensitive to whether the missing-CS2 worker-years are imputed or dropped. The college sub-panel does depend on the imputation, but only because the alternative is a sample that omits a relevant set of public-sector employees by construction; the differences in Table 25 reflect the non-representativeness of the no-imputation sample rather than any bias in the imputation procedure.

E Supplementary Material on the RC / RM Split

Section 4.3.1 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 $o \to$CS2 entry-rate breakdown by destination and the full five-state unconditional transition matrices.

E.1 Entry Rates by Destination

Table 26 reports the $o \to$CS2 entry rate (annual average $o \to$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, CS54) and blue-collar rm codes (CS62, CS67) experience sizeable entry-rate drops, ruling out a pure deindustrialisation reading of the entry-margin collapse.

E.2 Five-State Unconditional Transition Matrices

Tables 27 and 28 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 j at t and state k at $t{+}1$; the matrix sums to one. Row sums give the origin distribution (the population share in state j at t); column sums give the destination

Table 26: $o \rightarrow$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 $o \rightarrow$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 = acute polarization 1994–2006 (excluding expansion pairs 2001–02 and 2005–06); Post-crisis = 2009–2019.

distribution. Comparing the two panels makes the key rc-specific transitions visible as joint probabilities: the $o \rightarrow rc$ cell falls while the $rc \rightarrow rc$ diagonal rises, and the rm row compresses as the rm population itself shrinks. Four gross-flow patterns jump out when

Table 27: Unconditional transition matrix — pre-polarization 1984–1993. Values ×100.

From j \ To k	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

comparing the two sub-periods — patterns that the net-flow decomposition of Table 13 collapses into single signed numbers. First, the $o \rightarrow rc$ cell falls from 2.49 to 1.76 (a 29% drop), while the $rc \rightarrow o$ 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 $o \rightarrow rm$ cell falls even more steeply, from 1.55 to 1.06 (32%

Table 28: Unconditional transition matrix — polarization 1994–2021. Values ×100.

From j \ To k	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

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 $m \to m$ and $rc \to rc$ diagonal entries rise substantially (7.63 to 10.89 and 18.55 to 20.65 respectively), reflecting higher within-state persistence as occupation-specific tenure lengthens over the sample. Fourth, the $a \to a$ diagonal rises by two-thirds (6.58 to 10.89), driven mostly by the rising abstract population rather than by individual persistence changes: the a row sum rises from 9.12 to 12.77, a 40% increase in population share. 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.

F Linear Probability Models: Full Regression Tables

This appendix reports the linear probability model (LPM) estimates underlying the discussion in Section 5. For each origin state $j \in \{m, r, a, o\}$ and each destination $k \neq j$, the dependent variable is an indicator for the one-year transition $j \to k$. All regressions include age centred at 40 (entering linearly, $\text{age}_c = \text{age} - 40$), a female dummy, tenure in the current two-digit PCS code (for j-origin LPMs with $j \in \{m, r, a\}$) or duration of the ongoing non-employment spell (for o-origin LPMs), and year fixed effects. For o-origin transitions the sample is restricted to *re-entrants*, i.e. workers with at least one observed employment year prior to t in the panel; education is imputed at worker level via forward and backward fill of the EDP record before this restriction so that an education value observed during a worker's employed years can fill her o-origin records too. Specification S1 (baseline, no education) uses all 35 matched year-pairs 1984–2021 (excluding the 1989→1991 gap pair); specifications S2–S3 restrict to post-2002 observations where education is available. The baseline transition rate $\bar{y}$ reported at the foot of each table is the unconditional probability of the transition within the origin-state sample.

We report six headline tables in full: the three upgrading transitions ($r \to a$, $m \to a$, $m \to r$), one representative downgrading transition ($a \to r$), and the two non-employment entries ($o \to r$, $o \to a$). Full tables for the remaining six transitions ($r \to m$, $a \to m$, $m \to o$, $r \to o$, $a \to o$, $o \to m$) are available on request.

Table 29: LPM: $r \to a$

	(1) S1		(2) S2: +College		(3) S3: +Fem×Col	
	Coef	(SE)	Coef	(SE)	Coef	(SE)
Constant	$+0.0477^{***}$	(0.0009)	$+0.0425^{***}$	(0.0003)	$+0.0407^{***}$	(0.0003)
Age (centred)	$+0.0001^{***}$	(0.0000)	$+0.0003^{***}$	(0.0000)	$+0.0003^{***}$	(0.0000)
Female	-0.0159^{***}	(0.0001)	-0.0199^{***}	(0.0001)	-0.0165^{***}	(0.0001)
College			$+0.0625^{***}$	(0.0003)	$+0.0892^{***}$	(0.0007)
Female × Col.					-0.0412^{***}	(0.0007)
Tenure (occupation)	-0.0030^{***}	(0.0000)	-0.0026^{***}	(0.0000)	-0.0026^{***}	(0.0000)
Year FE	Yes		Yes		Yes	
N	10,177,803		8,350,450		8,350,450	
R^2	0.0080		0.0190		0.0200	
$\bar{y}$	0.0321		0.0308		0.0308	

Notes: HC3 robust standard errors. Tenure is years in the current two-digit PCS code at $t-1$. S1 uses the full 1984–2021 panel (excluding the 1989→1991 gap pair); S2 and S3 restrict to post-2002 observations with observed education (worker-level forward/backward fill). College coefficient in S3 is the slope for men; the female-college slope equals the Female coefficient plus the Female × College interaction. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 30: LPM: $m \to a$

	(1) S1		(2) S2: +College		(3) S3: +Fem×Col	
	Coef	(SE)	Coef	(SE)	Coef	(SE)
Constant	$+0.0165^{***}$	(0.0009)	$+0.0145^{***}$	(0.0003)	$+0.0139^{***}$	(0.0003)
Age (centred)	-0.0002^{***}	(0.0000)	$+0.0000^{**}$	(0.0000)	$+0.0000$	(0.0000)
Female	-0.0091^{***}	(0.0001)	-0.0097^{***}	(0.0001)	-0.0087^{***}	(0.0001)
College			$+0.0504^{***}$	(0.0006)	$+0.0633^{***}$	(0.0011)
Female × Col.					-0.0226^{***}	(0.0013)
Tenure (occupation)	-0.0014^{***}	(0.0000)	-0.0012^{***}	(0.0000)	-0.0011^{***}	(0.0000)
Year FE	Yes		Yes		Yes	
N	3,993,507		3,185,811		3,185,811	
R^2	0.0050		0.0152		0.0157	
$\bar{y}$	0.0116		0.0104		0.0104	

Notes: HC3 robust standard errors. Tenure is years in the current two-digit PCS code at $t-1$. S1 uses the full 1984–2021 panel (excluding the 1989→1991 gap pair); S2 and S3 restrict to post-2002 observations with observed education (worker-level forward/backward fill). College coefficient in S3 is the slope for men; the female-college slope equals the Female coefficient plus the Female × College interaction. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 31: LPM: $m \to r$

	(1) S1		(2) S2: +College		(3) S3: +Fem×Col	
	Coef	(SE)	Coef	(SE)	Coef	(SE)
Constant	$+0.2344^{***}$	(0.0036)	$+0.1668^{***}$	(0.0009)	$+0.1678^{***}$	(0.0009)
Age (centred)	-0.0029^{***}	(0.0000)	-0.0023^{***}	(0.0000)	-0.0023^{***}	(0.0000)
Female	-0.0278^{***}	(0.0003)	-0.0282^{***}	(0.0003)	-0.0300^{***}	(0.0003)
College			$+0.0577^{***}$	(0.0010)	$+0.0340^{***}$	(0.0015)
Female × Col.					$+0.0416^{***}$	(0.0020)
Tenure (occupation)	-0.0088^{***}	(0.0000)	-0.0080^{***}	(0.0000)	-0.0080^{***}	(0.0000)
Year FE	Yes		Yes		Yes	
N	3,993,507		3,185,811		3,185,811	
R^2	0.0473		0.0458		0.0460	
$\bar{y}$	0.1065		0.0945		0.0945	

Notes: HC3 robust standard errors. Tenure is years in the current two-digit PCS code at $t-1$. S1 uses the full 1984–2021 panel (excluding the 1989→1991 gap pair); S2 and S3 restrict to post-2002 observations with observed education (worker-level forward/backward fill). College coefficient in S3 is the slope for men; the female-college slope equals the Female coefficient plus the Female × College interaction. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 32: LPM: $a \to r$

	(1) S1		(2) S2: +College		(3) S3: +Fem×Col	
	Coef	(SE)	Coef	(SE)	Coef	(SE)
Constant	$+0.1927^{***}$	(0.0042)	$+0.1254^{***}$	(0.0009)	$+0.1243^{***}$	(0.0009)
Age (centred)	-0.0003^{***}	(0.0000)	-0.0009^{***}	(0.0000)	-0.0009^{***}	(0.0000)
Female	$+0.0086^{***}$	(0.0003)	$+0.0153^{***}$	(0.0003)	$+0.0195^{***}$	(0.0005)
College			-0.0504^{***}	(0.0003)	-0.0476^{***}	(0.0003)
Female × Col.					-0.0082^{***}	(0.0006)
Tenure (occupation)	-0.0068^{***}	(0.0000)	-0.0059^{***}	(0.0000)	-0.0059^{***}	(0.0000)
Year FE	Yes		Yes		Yes	
N	3,592,038		2,938,659		2,938,659	
R^2	0.0270		0.0352		0.0353	
$\bar{y}$	0.0635		0.0625		0.0625	

Notes: HC3 robust standard errors. Tenure is years in the current two-digit PCS code at $t-1$. S1 uses the full 1984–2021 panel (excluding the 1989→1991 gap pair); S2 and S3 restrict to post-2002 observations with observed education (worker-level forward/backward fill). College coefficient in S3 is the slope for men; the female-college slope equals the Female coefficient plus the Female × College interaction. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 33: LPM: $o \to r$ (re-entrants)

	(1) S1		(2) S2: +College		(3) S3: +Fem×Col	
	Coef	(SE)	Coef	(SE)	Coef	(SE)
Constant	$+0.6534^{***}$	(0.0061)	$+0.6219^{***}$	(0.0036)	$+0.6369^{***}$	(0.0036)
Age (centred)	-0.0023^{***}	(0.0000)	-0.0036^{***}	(0.0000)	-0.0035^{***}	(0.0000)
Female	$+0.0068^{***}$	(0.0008)	$+0.0047^{***}$	(0.0010)	-0.0212^{***}	(0.0010)
College			-0.1453^{***}	(0.0014)	-0.2421^{***}	(0.0020)
Female × Col.					$+0.1797^{***}$	(0.0027)
Tenure (NE spell)	-0.0075^{***}	(0.0002)	-0.0078^{***}	(0.0002)	-0.0079^{***}	(0.0002)
Year FE	Yes		Yes		Yes	
N	1,369,102		1,064,051		1,064,051	
R^2	0.0201		0.0296		0.0336	
$\bar{y}$	0.5217		0.5201		0.5201	

Notes: HC3 robust standard errors. Sample restricted to re-entrants (workers with at least one observed employment year prior to t). Tenure is the duration of the current non-employment spell ending at t. S1 uses the full 1984–2021 panel (excluding the 1989→1991 gap pair); S2 and S3 restrict to post-2002 observations with observed education (worker-level forward/backward fill). College coefficient in S3 is the slope for men; the female-college slope equals the Female coefficient plus the Female × College interaction. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 34: LPM: $o \to a$ (re-entrants)

	(1) S1		(2) S2: +College		(3) S3: +Fem×Col	
	Coef	(SE)	Coef	(SE)	Coef	(SE)
Constant	$+0.1891^{***}$	(0.0044)	$+0.1491^{***}$	(0.0024)	$+0.1417^{***}$	(0.0024)
Age (centred)	$+0.0009^{***}$	(0.0000)	$+0.0028^{***}$	(0.0000)	$+0.0028^{***}$	(0.0000)
Female	-0.0940^{***}	(0.0006)	-0.0985^{***}	(0.0006)	-0.0858^{***}	(0.0006)
College			$+0.3877^{***}$	(0.0013)	$+0.4353^{***}$	(0.0019)
Female × Col.					-0.0884^{***}	(0.0026)
Tenure (NE spell)	-0.0044^{***}	(0.0001)	-0.0029^{***}	(0.0001)	-0.0029^{***}	(0.0001)
Year FE	Yes		Yes		Yes	
N	1,369,102		1,064,051		1,064,051	
R^2	0.0348		0.1838		0.1857	
$\bar{y}$	0.1534		0.1470		0.1470	

Notes: HC3 robust standard errors. Sample restricted to re-entrants (workers with at least one observed employment year prior to t). Tenure is the duration of the current non-employment spell ending at t. S1 uses the full 1984–2021 panel (excluding the 1989→1991 gap pair); S2 and S3 restrict to post-2002 observations with observed education (worker-level forward/backward fill). College coefficient in S3 is the slope for men; the female-college slope equals the Female coefficient plus the Female × College interaction. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.