Quarterly Labor Input Measurement: Extending the Industry-Level Production Account to Business Cycle Frequencies

Authors David Jang and Jon D. Samuels, U.S. Bureau of Economic Analysis[1]

Author of Contact david.jang@bea.gov

Date September 2026

Abstract Industry-level labor input is conventionally measured annually, which averages away the sharpest movements in hours and workforce composition during rapid recessions. This paper extends the U.S. Bureau of Economic Analysis annual Industry-Level Production Account to a quarterly frequency for 63 industries over 2003–2024. To overcome the sparsity of Current Population Survey microdata, we combine state-space smoothing of marginals, Denton-Cholette benchmarking, and iterative proportional fitting to recover a reconciled distribution. The resulting quarterly quality-adjusted labor input series demonstrates that annual aggregation cannot resolve short-duration shocks. Because the COVID–19 contraction and early recovery occurred within a single calendar year, the peak-to-trough raw-hours decline measured quarterly is 6.7 percentage points deeper than its annual counterpart, and the labor composition spike is 1.3 points larger. Furthermore, quarterly data reveal that COVID–19's composition shocks were concentrated and transitory, whereas the Great Recession's were broad and persistent. At the industry level, the COVID–19 quarterly-annual measurement differences are near-uniform shifts across industries, and the industries with the largest composition spikes reverse the most through the recovery.

Keywords labor input, labor composition, quarterly measurement, business cycles, recovery dynamics

JEL codes C82, E01, E24, E32, J24

[1] We thank participants at the U.S. Bureau of Economic Analysis internal seminar and the 2026 Society of Economic Measurement conference for helpful discussions. The views expressed in this paper are those of the authors and do not necessarily represent the views of the Bureau of Economic Analysis or the U.S. Department of Commerce.

Suggested citation:
Jang, David and Jon D. Samuels. *Quarterly Labor Input Measurement: Extending the Industry-Level Production Account to Business Cycle Frequencies.* Working paper no. WP2026–19. Washington, DC: U.S. Bureau of Economic Analysis, September 2026, https://doi.org/10.66137/########.

1. Introduction

Recessions unfold within and across calendar years, but the U.S. Bureau of Economic Analysis Industry-Level Production Account (ILPA) currently provides annual labor input measures. Annual observations therefore average peaks, troughs, and early recoveries into a single value. This timing problem is especially important for short, concentrated shocks. The Great Recession (GFC in figure labels) contracted over six quarters spanning nearly three calendar years, whereas the COVID-19 contraction was concentrated in two quarters within 2020.

This paper extends the annual ILPA labor input framework to quarterly frequency while preserving its detailed worker-type-by-industry structure. We refer to the resulting quarterly quality-adjusted labor input series as "QALI." While ILPA refers to the broader industry-level production account framework in which labor input can be combined with quarterly capital and intermediate inputs, this paper focuses on the measurement of labor input itself and on how raw hours and labor composition move over the business cycle.

The paper makes three main contributions. First, we develop an integrated Kalman-Denton-IPF estimation pipeline, combining state-space smoothing, Denton-Cholette benchmarking, and multiway iterative proportional fitting (IPF), for constructing quarterly labor input from noisy Current Population Survey (CPS) microdata. Rather than directly estimating 12,000 detailed worker-type-by-industry cells, our sequential method smooths and benchmarks lower-dimensional marginal vectors before distributing those targets back to the detailed cells. Second, we demonstrate that the direction of the quarterly-annual measurement gap depends on the shape of the within-year path. Because the COVID-19 contraction and early recovery occurred within a single calendar year, the quarterly peak-to-trough measures of both the raw-hours drawdown and the labor composition spike are substantially larger than their annual counterparts. In contrast, annual data closely track the slow, multi-year progression of the Great Recession. Third, we characterize labor composition over the business cycle. We find that COVID-19's composition shocks were large, highly concentrated, and broadly transitory. The Great Recession's shifts, by contrast, were broader and persistent.

The measurement problem arises from the granularity required by the annual ILPA framework. Labor input is constructed for 192 structural worker types defined by education, age, sex, and class of worker within 63 industries, yielding approximately 12,000 structural cells.[1] Quarterly CPS samples are too sparse for reliable direct estimation at this resolution. The CPS samples about 60,000 households each month, roughly 110,000 to 130,000 persons, so the employed respondents in a quarter's three pooled monthly files average fewer than 20 per detailed cell. At the same time, collapsing the data would discard the compensation-share heterogeneity required for Törnqvist quality adjustment.

We resolve this tension in three modules. First, local-linear-trend state-space models learn trend innovation variances from annual ILPA marginal series and use those dynamics to smooth quarterly CPS marginal indicators. Second, Denton-Cholette benchmarking anchors the smoothed quarterly marginals to annual ILPA means. Employment is then proportionally aligned to aggregate quarterly CPS employment. This alignment moves the marginal-stage employment level away from the ILPA benchmark by a stable wedge of about -5 percent, a level difference between the CPS and ILPA employment data, quantified in Section 4.2. A final cell-level alignment after IPF restores annual ILPA consistency in the detailed data, so the CPS adjustment's lasting contribution is the within-year quarterly shape of employment rather than its level. Hours, weeks, and compensation take their within-year shape from the smoothing and benchmarking stages, and the CPS adjustment leaves that shape unchanged. Third, IPF uses interpolated annual ILPA cell distributions as seeds and reconciles them to the available quarterly marginal targets for employment (E) and the product quantities EH, EW, and $EHWC$, where H, W, and C denote hours per week, weeks per year, and hourly compensation, so that EH is aggregate weekly hours, EW aggregate annual weeks, and $EHWC$ total labor compensation. The set of binding marginal constraints differs by product, as detailed in Section 4.3.

The remainder of the paper proceeds as follows. Section 2 reviews the related labor-input and temporal-disaggregation literature. Section 3 describes the data and the ILPA labor-input framework. Section 4 presents the quarterly construction. Section 5 reports the recession hours and composition results. Section 6 concludes. The appendix documents variable definitions and implementation details.

[1] The youngest structural age group, ages 1-15, is retained for classification consistency but zeroed out before labor input aggregation, leaving 168 active worker types.

Figure 1 previews the finished data product. It plots the aggregate quarterly QALI series against its annual counterpart over 2003-2024, with the two recession windows shaded. The quarterly series stays anchored to the annual account through normal times, so the quarterly extension does not drift from the published benchmark, yet it resolves the sharp within-year movements that annual averaging cannot. The COVID-19 episode is the clearest case. Quarterly QALI falls to 109.5 at the 2020Q2 trough while the flat annual 2020 observation reads 114.9, a within-year gap of 5.4 index points that annual data average away. The Great Recession, spread across calendar years, shows the quarterly and annual paths tracking closely. This contrast is explained in detail in Section 5.

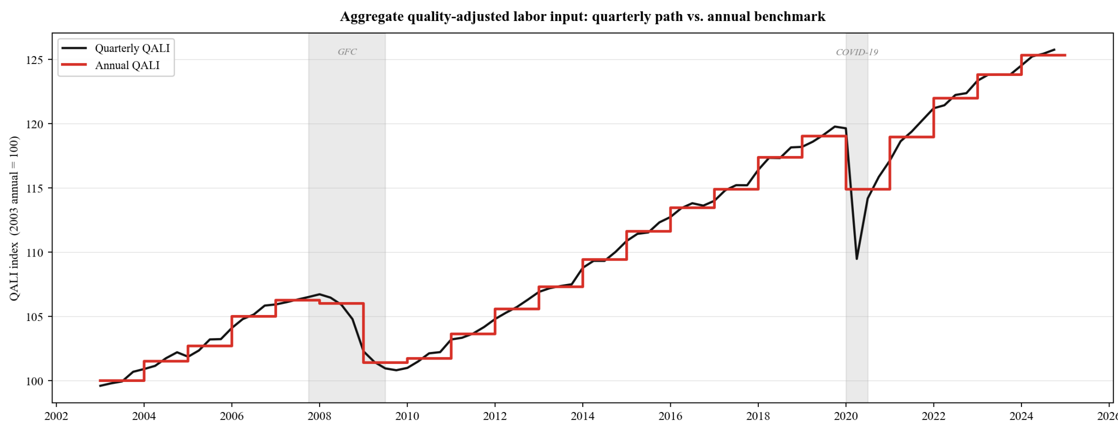

Figure 1. **Aggregate quality-adjusted labor input (QALI), quarterly path versus annual benchmark, 2003-2024, indexed to the 2003 annual mean. The black line is the quarterly series and the red steps are the annual series, flat within each calendar year in the same format as Figures 3 and 4. Both are the compensation-share-weighted two-stage Törnqvist aggregate of the 63 industries' labor input. Shaded bands mark the Great Recession and COVID-19. The quarterly series threads through the annual benchmarks in normal times and departs from them within the calendar year during the COVID-19 contraction.**

2. Literature Review

This paper contributes to three strands of literature, namely labor quality measurement in production accounting, the macro-labor literature on cyclical composition bias, and temporal disaggregation.

The measurement of labor quality in growth accounting descends from Jorgenson and

Griliches (1967), who showed that aggregating hours without accounting for worker heterogeneity mismeasures labor input growth, and from Jorgenson et al. (1987), who operationalized compensation-share-weighted labor input for U.S. industries. The ILPA framework, jointly developed by the U.S. Bureau of Economic Analysis (BEA) and the U.S. Bureau of Labor Statistics (BLS), embeds this approach in a consistent industry production account, decomposing labor input growth into raw hours and labor composition. Garner et al. (2018) provides the foundational methodology for the 63-industry framework adopted here. Stewart (2022) shows that disproportionate employment losses among lower-wage workers materially raised measured labor quality during COVID-19. We extend this line of work by measuring the same worker-composition mechanism quarterly, allowing the timing of employment losses and composition shifts to be evaluated within annual benchmark years.

A parallel macro-labor literature establishes that the composition of the employed workforce varies systematically over the business cycle and biases cyclical readings of aggregate averages. Bils (1985) documents that aggregate real wages understate the procyclicality visible in panel data on individual workers, and Solon et al. (1994) show that this gap arises from composition bias. Low-wage workers' hours are disproportionately cyclical, so the composition of hours shifts toward high-wage workers in downturns. Our composition results are the industry-level, production-account counterpart of that mechanism. The same disproportionate loss of low-wage hours that biases aggregate wage cyclicality appears here as a countercyclical labor-quality wedge, measured quarterly for each of 63 industries within a Törnqvist accounting structure.

The closest existing data products each combine two of the three properties that define our measure (quality adjustment, industry detail, and quarterly frequency) but not all three. The annual ILPA account (Garner et al., 2018) is industry-resolved and quality-adjusted but annual, and our series is its quarterly extension. Fernald (2014) maintains a widely used quarterly total factor productivity (TFP) series whose labor-quality component is estimated from CPS demographic shares, but it covers only the aggregate business sector. The Federal Reserve Bank of Chicago's quarterly industry-level labor productivity dataset of Hobijn et al. (2025) provides timely output-per-hour measures for 87 industries, but its labor input is unadjusted hours consistent with the BLS Productivity and Costs measure. The 63-industry quarterly QALI combines all three properties. It preserves the worker-type-by-industry Törnqvist structure of the annual account and benchmarks to its annual means, so *industry-level* composition, the margin on which our recession results turn, remains tied to the published annual ILPA.

The temporal-disaggregation literature provides the foundation for constructing quarterly paths from annual benchmarks. Denton (1971) introduced a quadratic-minimization approach that preserves movement in a high-frequency indicator subject to low-frequency constraints. The Denton-Cholette variant (Dagum and Cholette, 2006) corrects the boundary distortions of the original formulation. Regression-based alternatives such as Chow and Lin (1971), and more recent state-space and dynamic-factor interpolators, could in principle be applied series by series. We instead combine state-space smoothing of noisy CPS indicators, Denton-Cholette benchmarking, and IPF because the object of interest is not a single series but a mutually consistent *system* of five overlapping marginal vectors and roughly 12,000 detailed cells linked by accounting identities. Joint model-based estimation at that dimension is computationally impractical, and the IPF stage is what enforces the cross-sectional consistency that per-series interpolators cannot deliver.

3. Data

We utilize Integrated Public Use Microdata Series Current Population Survey data (IPUMS CPS; Flood et al., 2025) to construct the base distributions of workers across worker-type and industry cells. The basic monthly CPS data provide individual-level microdata on worker characteristics (age, sex, and education), employment status, industry classification, and hours worked. The CPS outgoing rotation group (ORG) data provide wage/income and weeks worked per year information. We use data for the period 2003-2024. The industry classification is consolidated to 63 North American Industry Classification System (NAICS) groups and worker-type variables are harmonized to the ILPA definitions documented in Appendix A.

The ILPA provides annual benchmarks for employment (E), hours per week (H), weeks worked per year (W), and hourly compensation (C). For benchmarking and cell reconciliation, we work with the additive product quantities E, EH, EW, and $EHWC$, that is, employment, aggregate weekly hours, aggregate annual weeks, and total labor compensation, each formed as the sum of the corresponding cell-level products. Unlike the per-worker rates H, W, and C, the products aggregate additively across cells, which is what the Denton-Cholette and IPF stages require. The benchmarks are available by detailed NAICS industry classification. Garner et al. (2018) provides the methodology and implementation of the integrated account, which combines BEA's industry output and intermediate input estimates, including capital input, with BLS labor input data to produce a consistent production account framework.

Our quarterly disaggregation builds directly on the ILPA annual framework, using the published estimates as annual-mean constraints in the Denton-Cholette stage. An intermediate step aligns quarterly employment with the aggregate CPS series, which temporarily moves employment away from the annual ILPA benchmark, so Section 4.2 documents that adjustment separately. The final detailed cells are realigned to the ILPA annual benchmarks after IPF (Section 4.4).[2]

Labor composition measurement uses the Törnqvist index to decompose labor input growth into hours and quality components. At the cell level (worker-type group g within industry i), we observe employment $E_{g,i}$, hours per worker $H_{g,i}$, weeks per year $W_{g,i}$, and hourly compensation $C_{g,i}$. Total hours in cell (g,i) are $\mathsf{HRS}_{g,i,t} = E_{g,i,t} \times H_{g,i,t} \times W_{g,i,t}$.

We compute these growth measures separately for each of the 63 BEA industries. Quality-adjusted labor growth (composition-adjusted hours) uses the Törnqvist index, weighting each worker-type group's hours growth by its cost share:

$$\mathsf{QLG}_{i,t} = \sum_{g} \frac{1}{2}(\hat{s}_{g,i,t} + \hat{s}_{g,i,t-1}) \ln\left(\frac{\mathsf{HRS}_{g,i,t}}{\mathsf{HRS}_{g,i,t-1}}\right), \tag{1}$$

where $\hat{s}_{g,i,t} = \frac{C_{g,i,t}\mathsf{HRS}_{g,i,t}}{\sum_g C_{g,i,t}\mathsf{HRS}_{g,i,t}}$ is the compensation-weighted hour share of group g in industry i, $C_{g,i,t}$ is the cell's hourly compensation, and weights are averaged across two periods.

[2] The ILPA data also provide the reference weights for separating the NAICS 525 industry in CPS into industries 42 and 44. This is necessary since the CPS data does not distinguish these two industries.

Raw labor (hours) growth is:

$$\mathsf{LG}_{i,t} = \ln\left(\frac{\sum_g \mathsf{HRS}_{g,i,t}}{\sum_g \mathsf{HRS}_{g,i,t-1}}\right). \tag{2}$$

The quality (composition) growth component is the difference:

$$\mathsf{QG}_{i,t} = \mathsf{QLG}_{i,t} - \mathsf{LG}_{i,t}. \tag{3}$$

This difference captures shifts in the composition of hours toward workers in groups with higher average labor compensation (positive growth rates) or lower average labor compensation (negative growth rates), reflecting labor quality changes from compositional shifts.

The first step of the data integration process harmonizes CPS worker characteristics to the worker-type dimensions used in ILPA. These dimensions are crossed to form the worker-type cells used in the Törnqvist labor input calculation. Table 1 lists the dimensions and their categories.

Table 1. Worker-Type Dimensions for Labor Input Measurement

Dimension	Number of Groups	Categories
Age	8	1-15; 16-17; 18-24; 25-34; 35-44; 45-54; 55-64; 65+
Education	6	Less than 9th grade; 9th-12th grade, no diploma; high school diploma/GED; some college or associate degree; bachelor's degree; graduate degree (master's, professional, or doctoral)
Sex	2	Male; female
Class of worker	2	Wage and salary (private and government); self-employed

Crossing these dimensions yields $8 \times 6 \times 2 \times 2 = 192$ structural worker-type cells. The ages 1-15 group is retained for classification consistency but zeroed out before labor input aggregation, so the active worker-type universe contains 168 cells.

Additionally, 63 industries are mapped to 13 sectors in order to ensure that the worker-type and industry cross-classification has enough observations when building marginal vectors. The 13 broad sectors include (1) agriculture, (2) mining and utilities, (3) construction, (4) durable manufacturing, (5) nondurable manufacturing, (6) trade, (7) transportation, (8) information and financial services, (9) professional and business services, (10) education and health, (11) leisure and hospitality, (12) other services, and (13) government. The full industry-to-sector mapping is listed in Appendix B.1.

4. Methodology Overview

Quarterly CPS microdata are too noisy for direct estimation of the full worker-type-by-industry array. The quarterly construction therefore operates first on five lower-dimensional marginal vectors, namely worker type (education $\times$ age $\times$ sex), industry, class of worker $\times$ sector, education $\times$ sector, and sex $\times$ sector. The method has three modules, followed by a final annual alignment of the detailed cells. The construction follows the same basic approach as the annual ILPA account (Garner et al., 2018) and the growth-accounting framework of Jorgenson et al. (2005).

The estimation pipeline is sequential, as each of the three modules addresses a distinct deficiency in the high-frequency microdata. First, because quarterly CPS marginals are dominated by sampling noise, they are smoothed into credible high-frequency indicators. Second, because smoothing alone cannot guarantee correct aggregate levels, these indicators are benchmarked to the annual ILPA account, pinning down the structural level while preserving the high-frequency shape. Finally, only after the marginal totals are corrected for both noise and level can they be distributed down to the roughly 12,000 detailed cells using IPF. A fourth and final step, applied after IPF, aligns each detailed cell's annual mean to its ILPA cell benchmark (Section 4.4).

4.1. State-space smoothing of quarterly marginals

Each quarterly CPS marginal series is first seasonally adjusted using X-13ARIMA-SEATS, with implementation details in Appendix B.2. For each annual ILPA marginal series, we estimate a local linear trend model:

$$\ln(X_y^{\text{annual}}) = \mu_y + \epsilon_y, \tag{4}$$

$$\mu_{y+1} = \mu_y + \beta_y + \eta_{\mu,y}, \tag{5}$$

$$\beta_{y+1} = \beta_y + \eta_{\beta,y}, \tag{6}$$

where μ_y is the level, β_y is the slope, and the disturbances have variances R, Q_μ, and Q_β. Annual equations are indexed by y and quarterly equations by q. The annual series identify the process innovation variances Q_μ and Q_β. These dynamics are then applied to the corresponding quarterly CPS marginal indicator, while quarterly measurement variance is estimated separately to reflect higher survey noise. The smoothing is performed on marginal series, not on the approximately 12,000 detailed cells.

One deliberate exception applies. Employment for the worker-type marginal (education $\times$ age $\times$ sex) is not Kalman-smoothed, and the seasonally adjusted series is used directly as the benchmarking indicator. The purpose of the exception is to keep aggregate labor-composition change timely. The worker-type employment shares are the direct source of that change, and the filter would delay and damp exactly those movements. Bypassing the filter is acceptable for this series because employment has the largest sample sizes of the four variables, so its quarterly movements are the most reliably measured. The sector-crossed marginals and all rate variables retain the full smoothing treatment. Hours, weeks, and compensation for the worker-type marginal remain smoothed. Because they are recovered from benchmarked product ratios, the unsmoothed employment factor largely cancels and their smoothness is preserved.

4.2. Denton-Cholette benchmarking and CPS employment alignment

Let I_{yq} denote the quarterly marginal indicator produced by the smoothing stage, the Kalman-smoothed series in general and the seasonally adjusted series for worker-type employment, and let x_{yq} denote its benchmarked value. Denton-Cholette preserves the within-year movement of I_{yq} subject to the annual-mean constraint

$$\frac{1}{4}\sum_{q=1}^{4} x_{yq} = X_y^{\text{ILPA}}. \tag{7}$$

The procedure is applied to the additive quantities E, EH, EW, and $EHWC$.

Figure 2 illustrates the smoothing and benchmarking stages for one industry's employment. The raw quarterly CPS tabulation is too noisy to use directly, with a quarter-to-quarter standard deviation about three times that of the smoothed path. The Kalman step extracts the underlying trend, and the Denton-Cholette step locks the annual mean of the smoothed indicator onto the ILPA benchmark while preserving its within-year movement. The level shift between the smoothed CPS indicator and the benchmarked path reflects the difference between the CPS employment level and the ILPA benchmark for this industry.

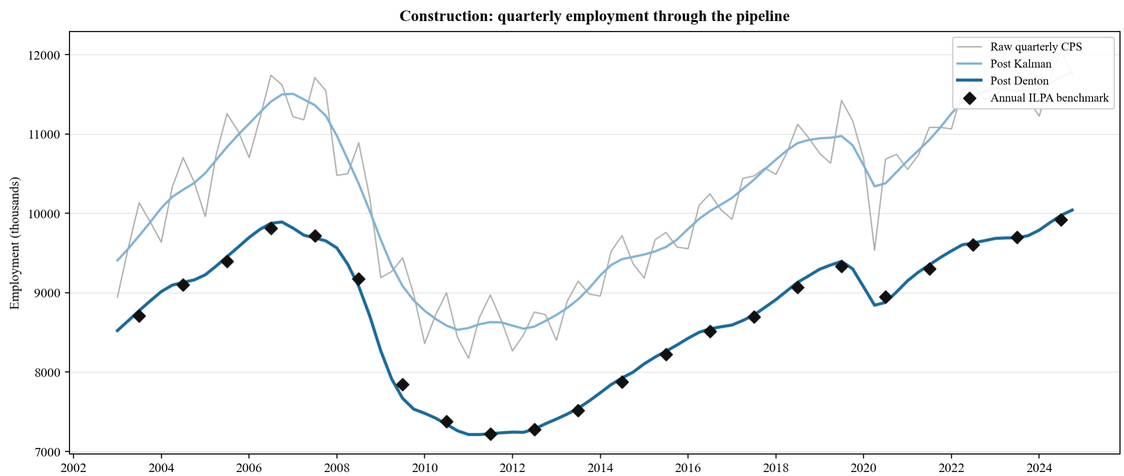

Figure 2. **The smoothing and benchmarking stages applied to construction employment, one of the 63 industry marginals. The grey line is the raw quarterly CPS tabulation, the light line the Kalman-smoothed indicator, and the dark line the Denton-Cholette benchmarked path. The diamonds are the annual ILPA employment benchmarks, which the benchmarked path reproduces in annual mean. Construction employment is Kalman-smoothed under the kept configuration, so the figure is representative of the general treatment rather than the worker-type employment exception.**

A second calibration aligns aggregate quarterly employment with seasonally adjusted CPS employment. This stage operates on the marginal vectors, since the detailed worker-type-by-industry array is not constructed until the IPF module. If $\widetilde{E}_{m,q}$ is the Denton-benchmarked employment of cell m within marginal vector s, the quarterly factor, computed and applied within each marginal vector, is

$$\lambda_{s,q} = \frac{E^{\mathsf{CPS}}_{s,q}}{\sum_{m \in s} \widetilde{E}_{m,q}}, \qquad E_{m,q} = \lambda_{s,q} \widetilde{E}_{m,q}, \tag{8}$$

where $E_{s,q}^{\text{CPS}}$ is the seasonally adjusted CPS employment aggregate for vector s, computed within the pipeline as the sum of the vector's X-13-adjusted cell indicators. This step preserves the quarterly cross-sectional distribution but moves the marginal-stage employment level away from the annual ILPA benchmark. The resulting wedge is quantifiable and stable. The implied annual mean of the CPS-aligned marginal employment sits below the ILPA benchmark by 5.3 percent on average over 2003-2024, within a narrow -4.3 to -6.4 percent band, reflecting the data difference between the CPS and the ILPA. Because the wedge is nearly constant, the within-year shape this stage imposes is essentially unaffected by it, as the year-to-year change in the wedge averages 0.3 percentage point. The wedge does not persist into the final data. The annual cell-level alignment of Section 4.4 subsequently rescales the detailed cells to the ILPA annual benchmarks, so the Stage 2 adjustment's lasting contribution is the quarterly shape of employment rather than its level. Hours, weeks, and hourly compensation retain their benchmarked paths and enter the product identities with the adjusted employment component.

4.3. Iterative proportional fitting

The third module starts from an interpolated annual ILPA worker-type-by-industry seed and reconciles it to the quarterly marginals using multiway IPF. Ordinary row-column RAS is the two-dimensional special case. The implementation here cycles over several overlapping marginal vectors in a five-dimensional array.

The reconciliation proceeds sequentially through the additive products, first E, then EH, then EW, then $EHWC$, followed by recovery of the component variables:

$$H = \frac{EH}{E}, \qquad W = \frac{EW}{E}, \qquad C = \frac{EHWC}{E \times H \times W}. \tag{9}$$

Before IPF, the five marginal vectors are harmonized so that mutually consistent targets exist. Employment totals of the industry, class-by-sector, education-by-sector, and sex-by-sector vectors are rescaled each quarter to the worker-type vector's total, which serves as the authoritative aggregate. The education-by-sector and sex-by-sector targets are then reconciled by two-dimensional IPF to education and sex controls derived from the worker-type vector and sector controls derived from the industry vector, and the class-by-sector targets are scaled to the same sector controls. This harmonization applies to all four products. Without it the five separately smoothed and benchmarked vectors would give the IPF loop mutually inconsistent targets. Appendix B.4 gives details.

For employment and employment-hours, all five marginal vectors are iteratively enforced. For EW and $EHWC$, the iterative loop binds only the worker-type, education-by-sector, and sex-by-sector constraints. Industry and class-by-sector totals are pre-reconciled but are not iteratively imposed because those CPS cells are especially sparse.

4.4. Annual cell-level alignment

After IPF, a final calibration aligns the detailed cells to the annual account. For each detailed worker-type-by-industry cell and year, the four quarterly values are averaged and compared with the corresponding annual ILPA cell benchmark. The cell's four quarters are then rescaled by the single factor that equates the two, with conservative safeguards for empty cells (Appendix B.5). Because one factor applies to all four quarters of a cell-year, the within-year movement produced by the earlier stages is preserved exactly while annual ILPA consistency is restored at the cell level. All labor-composition results in this paper are computed from the aligned cells.

5. Empirical Findings

The empirical analysis uses the quarterly labor input series to characterize how raw hours and labor composition move over the business cycle, and how much quarterly resolution matters relative to annual data. It is organized as a two-by-two. The aggregate level comes first, covering the peak-to-trough (Section 5.1) and then the recovery (Section 5.2). The industry cross-section follows for the same two phases (Sections 5.3 and 5.4). Two findings recur. At the trough, annual averaging attenuates whatever moves sharply within the calendar year. It understates both the COVID-19 raw-hours drawdown and the COVID-19 composition spike, while modestly overstating the Great Recession's slower-building composition change. Through the recovery, both margins move within the calendar year in ways annual data cannot resolve, and the two recessions differ in persistence rather than mechanism.

5.1. Peak-to-trough measurement: aggregate hours and composition

We first measure how much annual averaging distorts the peak-to-trough movement of the two labor margins at the aggregate level. Aggregate raw hours are total hours summed across the 63 industries, and their change is the log change of this total. Aggregate labor composition is the compensation-share-weighted two-stage Törnqvist index of labor input, QALI, minus the raw hours aggregate. It splits additively into a within-industry component (worker-type upgrading within industries) and a between-industry component (reallocation of hours across industries of differing pay). The quarterly measure uses common National Bureau of Economic Research (NBER) windows, 2007Q4-2009Q2 for the Great Recession and 2019Q4-2020Q2 for COVID-19. The annual measure applies the same aggregation to the annual ILPA benchmark series over the corresponding calendar years. All changes are peak-to-trough log-differences $\times 100$. Table 2 reports both margins.

Table 2. Aggregate Peak-to-Trough Raw Hours and Labor Composition

	Raw hours			Composition			Comp. mix	
Episode	Qtr	Ann	gap	Qtr	Ann	gap	within	between
Great Recession	-6.84	-6.86	$+0.02$	1.98	2.18	-0.20	1.15	0.83
COVID–19	-12.78	-6.08	-6.70	3.79	2.53	$+1.26$	2.81	0.98

Note: Peak-to-trough log-differences $\times 100$ (percentage points). Windows are 2007Q4 to 2009Q2 (Great Recession) and 2019Q4 to 2020Q2 (COVID-19). Annual entries compare the corresponding calendar years. Raw hours are the log change of total hours summed across the 63 industries. Composition is the compensation-share-weighted aggregate labor input (QALI) minus hours, split into within-industry worker-type composition and between-industry hours reallocation (within and between are the quarterly total). A negative hours gap means the quarterly drawdown is deeper. A negative composition gap means annual reads a larger composition change. Gap entries are computed from unrounded values and may differ in the final digit from the difference of the displayed columns.

Raw hours give the paper's central measurement result. In the Great Recession, the quarterly and annual drawdowns are essentially identical, because the contraction spans nearly three calendar years and annual averaging loses almost nothing. In COVID-19, however, the quarterly drawdown is more than twice as deep as the annual measure. Because the 2020Q2 trough and its rapid rebound fall entirely inside one calendar year, the annual observation averages the contraction, trough, and recovery into a single, much shallower value.

Labor composition rises in both episodes, but the COVID-19 spike is nearly twice as large, consistent with an hours drawdown about twice as deep concentrated in a much shorter window. The underlying mechanism is the same in both, as the disproportionate loss of low-wage hours mechanically raises measured average labor quality, but the composition mix differs. The Great Recession's rise is more evenly split between within-industry upgrading and between-industry reallocation. In contrast, the COVID-19 composition spike is dominated by the within-industry margin, as low-wage workers were shed within virtually every industry at once.

Comparing these quarterly readings to the annual data shows that the quarterly-annual gap depends on the duration of the shock relative to the calendar year. Because Great Recession composition builds slowly and keeps rising past the trough, the annual trough-year average sits slightly above the quarterly trough reading. In COVID-19, however, the composition spike peaks at the 2020Q2 trough and reverts substantially within the same calendar year, meaning annual averaging heavily attenuates the composition spike just as it attenuates the hours trough. Annual measurement thus understates both margins of the COVID-19 shock, while for the slow-moving Great Recession it tracks hours closely and modestly overstates composition. Section 5.3 traces this contrast to the industry cross-section.

5.2. Recovery dynamics: aggregate hours and composition

Because the quarterly series resolves within-year movement, it can trace how both margins evolve after the trough, not just their peak-to-trough change. For raw hours this is the mechanism behind the attenuation in Table 2. For composition it reveals a within-year cycle that annual data cannot see.

Raw hours trace opposite recovery shapes in the two episodes. Figure 3 plots the aggregate hours path relative to the prerecession baseline, with the annual series overlaid as a flat step within each calendar year. In COVID-19 the path is a sharp V. Hours fall to -12.8 log points at the 2020Q2 trough and rebound almost as fast, returning to the 2019Q4 baseline within about two years and rising above it thereafter. The flat annual 2020 observation sits near the midpoint of this V, at -6.1, which is why annual data understate both the drop and the rebound and why the hours drawdown gap in Table 2 is 6.7 points. The

Great Recession is instead a slow U in which hours remain depressed for years. Aggregate hours keep falling for two quarters past the NBER trough, reaching -8.1 in 2009Q4, and remain about 2.3 log points below baseline four years after the trough. There the annual steps track the quarterly path closely, consistent with the near-zero gap.

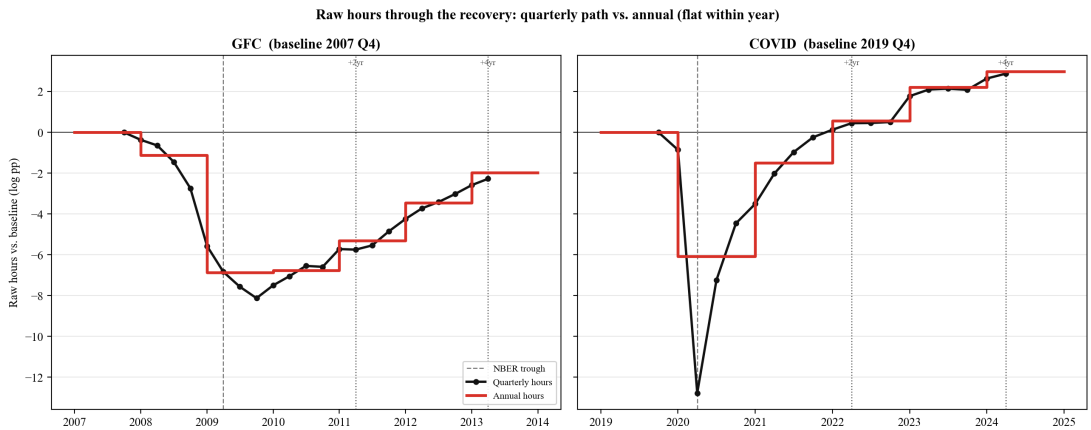

Figure 3. **Raw hours through the recovery, relative to the pre-recession baseline (2007Q4 for the Great Recession, 2019Q4 for COVID-19). The black line is the quarterly path of total hours. The red steps are the annual series, flat within each calendar year. The dashed vertical line marks the NBER trough, and the dotted lines mark two and four years after it. COVID-19 traces a sharp V that returns to baseline within about two years, with the flat annual 2020 observation near the midpoint. The Great Recession traces a slow U that annual data track closely.**

Labor composition resolves into a within-year cycle instead. Figure 4 plots the aggregate composition path relative to the prerecession baseline, through four years past the trough, with the annual series overlaid as a flat step within each calendar year, in the same format as Figure 3.

The two recoveries differ in persistence. In COVID-19, composition spikes at the trough ($+3.79$) and then reverses almost completely within about two years, falling to $+0.67$ by late 2021, before a smaller second rise through 2022-2024. In the Great Recession the rise is persistent. Composition builds to $+1.98$ at the trough, reaches $+2.63$ two quarters later, and stands at $+2.92$ four years after the trough.

The nonmonotonic COVID-19 path (a sharp spike, near-complete reversal, and later rerise) is invisible at annual frequency, where 2020 registers as a single elevated observation and

2021 as a single lower one. This within-year composition cycle is the clearest illustration of why quarterly resolution matters for measuring labor input over the business cycle.

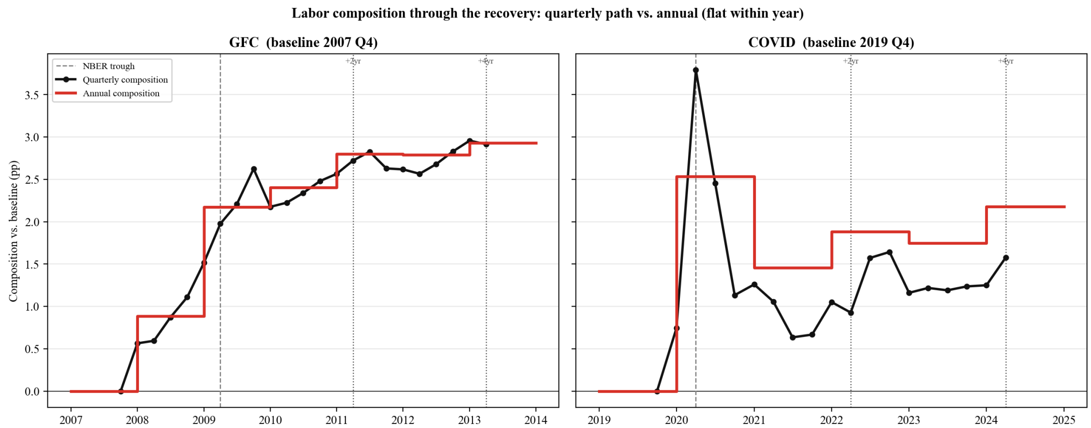

Figure 4. Labor composition through the recovery, relative to the pre-recession baseline (2007Q4 for the Great Recession, 2019Q4 for COVID-19), through four years past the trough. The black line is the quarterly composition path. The red steps are the annual series, flat within each calendar year. The dashed vertical line marks the NBER trough, and the dotted lines mark two and four years after it. The COVID-19 composition spike peaks at the trough and reverts within the calendar year, so the flat annual 2020 observation sits well below the quarterly spike. The Great Recession's composition builds slowly and keeps rising past the trough, and the annual steps track it closely.

5.3. Peak-to-trough measurement: the industry cross-section

The aggregate results show that annual averaging attenuates whatever reverts within the calendar year, which in COVID-19 means both the hours trough and the composition spike. The industry level reveals the underlying pattern, and shows that the gaps are near-uniform shifts in COVID-19 but small and idiosyncratic in the Great Recession. Figure 5 plots, for each of the 63 industries, the quarterly peak-to-trough change against its annual counterpart, separately for raw hours and labor composition in each recession. Points on the 45-degree line are measured identically at the two frequencies. Points off it are where annual measurement disagrees. Each panel also reports the mean quarterly-minus-annual gap.

For raw hours in COVID-19, the industry points form a tight band below the diagonal. Every industry's quarterly drawdown is deeper than its annual drawdown by a similar amount, 5.7 percentage points on average. Annual averaging thus understates the COVID-

19 hours drawdown almost uniformly, which is why the industry gaps aggregate to the -6.70 point gap in Table 2 and why the relative severity of the hardest-hit industries is essentially unchanged between the two frequencies. The largest individual gaps are not the hardest-hit industries but the pandemic *winners*, warehousing, transportation support, and data processing, whose within-year surge lifts the annual 2020 average above the 2020Q2 trough and places them above the band. In the Great Recession the hours points lie on the diagonal (mean gap $+0.17$), since that contraction spans three calendar years and annual data lose almost nothing.

Labor composition splits by episode. In the Great Recession the composition points fan closely around the diagonal with a small negative mean gap (-0.15). Annual measurement slightly overstates the composition rise there, and the disagreement is industry-specific rather than systematic, concentrated in a few small industries with sharp within-year swings (data processing, textile mills, apparel). In COVID-19 the picture mirrors the hours panel. The points sit *above* the diagonal in nearly every industry, with a mean gap of $+1.53$. Quarterly measurement reads a larger composition increase than annual almost uniformly, because industry composition spiked at the 2020Q2 trough and reverted within the calendar year, so the annual 2020 average sits well below the quarterly spike, the same attenuation mechanism that governs hours. The industries where the two frequencies disagree most in COVID-19 are computer and electronic products, legal services, and ambulatory health care.

Taken together, the industry cross-section shows that the quarterly-annual gap is episode-specific, not margin-specific. In COVID-19, annual data understate both the hours drawdown and the composition spike through near-uniform shifts that leave relative industry severity intact. In the Great Recession, annual data track hours closely and only slightly overstate composition, through a small, idiosyncratic gap that reflects composition's continued rise past the trough.

5.4. Recovery dynamics: the industry cross-section

The aggregate recovery path of Section 5.2 masks substantial cross-industry heterogeneity, and the industry level shows which industries drive the aggregate gaps and whether their composition changes reversed. Figure 6 plots, for each industry, the quarterly change

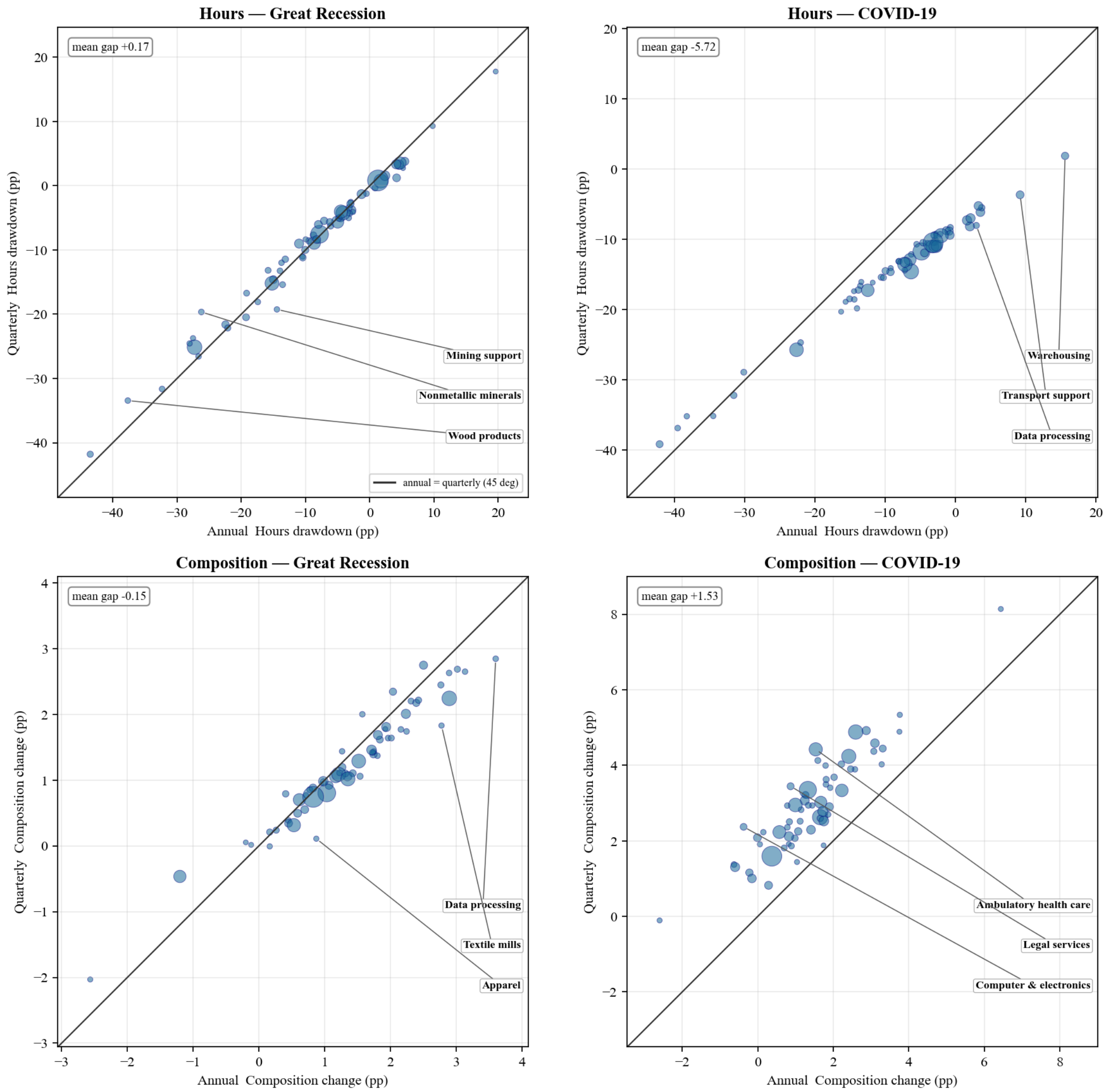

Figure 5. Quarterly versus annual peak-to-trough change by industry, for raw hours (top row) and labor composition (bottom row), in the Great Recession and COVID-19. Each point is one of the 63 industries (finance industries 42 and 44 dropped); marker size is the industry's hours share. The 45-degree line marks equality of the two frequencies. Each panel reports the mean quarterly-minus-annual gap. In COVID-19, hours form a tight band below the line and composition a band above it, both near-uniform shifts. In the Great Recession, both margins lie close to the line, with composition fanning idiosyncratically around it. The three labeled industries per panel have the largest absolute quarterly-minus-annual gap.

over the two years (eight quarters) following the trough against its annual counterpart (trough year to two years later), for raw hours and within-industry composition, in the same format as Figure 5. In COVID-19 the two margins deviate from the 45-degree line in opposite directions. The hours points lie above the line (mean gap $+6.48$ percentage points), as annual data understate the rebound just as they understate the drawdown, and the composition points lie below it (mean gap -2.05), as annual data miss the within-year reversal of the spike. In the Great Recession both margins lie close to the line (-0.63 for hours and $+0.15$ for composition).

On the hours margin the cross-section largely reproduces the aggregate paths of Section 5.2. In COVID-19 the industries with the deepest drawdowns rebound the most, with a cross-industry correlation of -0.48 between the peak-to-trough drawdown and the two-year recovery and a mean recovery of $+15.6$ points against a mean drawdown of -14.2. In the Great Recession hours barely recover across nearly the whole cross-section, with a mean two-year recovery of $+0.8$ points against a mean drawdown of -9.2 and a correlation of only -0.12. The composition margin is where the cross-section is more informative.

The two recessions differ sharply. In COVID-19, the industries with the largest composition spikes reverse the most. The cross-industry correlation between the trough spike and the composition change over the following eight quarters is -0.50, and the shock is concentrated in a few hard-hit industries (apparel, textiles, food manufacturing, administrative and support services) with large, transitory swings. Apparel alone spikes $+8.1$ and then reverses -7.4 percentage points. In the Great Recession the same correlation is $+0.53$. The industries whose composition rose most at the trough continued rising afterward, and the changes were smaller and more broadly spread across manufacturing and business services. Measured at four years rather than two, the contrast holds (-0.43 in COVID-19 and $+0.09$ in the Great Recession). The trough spike is also more dispersed across industries in COVID-19 (cross-industry standard deviation 1.31 versus 0.88). Finally, the industries driving the two episodes are largely different, with a cross-recession correlation of industry composition responses of only $+0.21$.

Taken together with the aggregate paths, the recovery evidence indicates that the two recessions' composition responses differed less in mechanism than in persistence. COVID-

19's were large, concentrated, and transitory, while the Great Recession's were smaller, broader, and persistent.

6. Conclusion

This paper develops a quarterly extension of the annual ILPA labor-input framework. The integrated Kalman-Denton-IPF pipeline smooths noisy quarterly CPS marginal indicators using dynamics learned from annual ILPA, benchmarks the quarterly paths to annual ILPA means, aligns employment with aggregate quarterly CPS employment, reconstructs the detailed worker-type-by-industry distribution while preserving the multiplicative $E \times H \times W \times C$ accounting structure, and finally aligns each detailed cell to its annual ILPA benchmark so that the published annual account is reproduced at the cell level.

The resulting quarterly series demonstrate why frequency matters. Because annual data mechanically absorb the rebounds that follow rapid troughs, they truncate our view of short-duration shocks. In COVID-19, the raw-hours drawdown and the labor-composition spike measured quarterly exceed their annual counterparts by economically meaningful margins. Furthermore, quarterly QALI resolves the dynamic path of labor composition over the cycle. While measured labor composition rose during both the Great Recession and COVID-19 as lower-paid workers lost hours disproportionately, the quarterly data reveal that the COVID-19 shock was concentrated and transitory, reversing almost completely within two years, whereas the Great Recession's composition shifts were smaller, broader, and persistent.

The contribution is both methodological and empirical. It comprises a quarterly labor-input data product that remains tied to the annual production-account structure, together with evidence that within-year timing materially affects both measured recession severity and the persistence of labor-composition changes that annual data cannot resolve.

The most immediate application is quarterly productivity measurement. Combining quarterly QALI with quarterly industry output would extend the production account to business-cycle frequency, and the composition results suggest the payoff is more than finer resolution. Because the 2020 composition spike was large and largely transitory, part of the measured surge in output per hour during the pandemic, and part of the subsequent

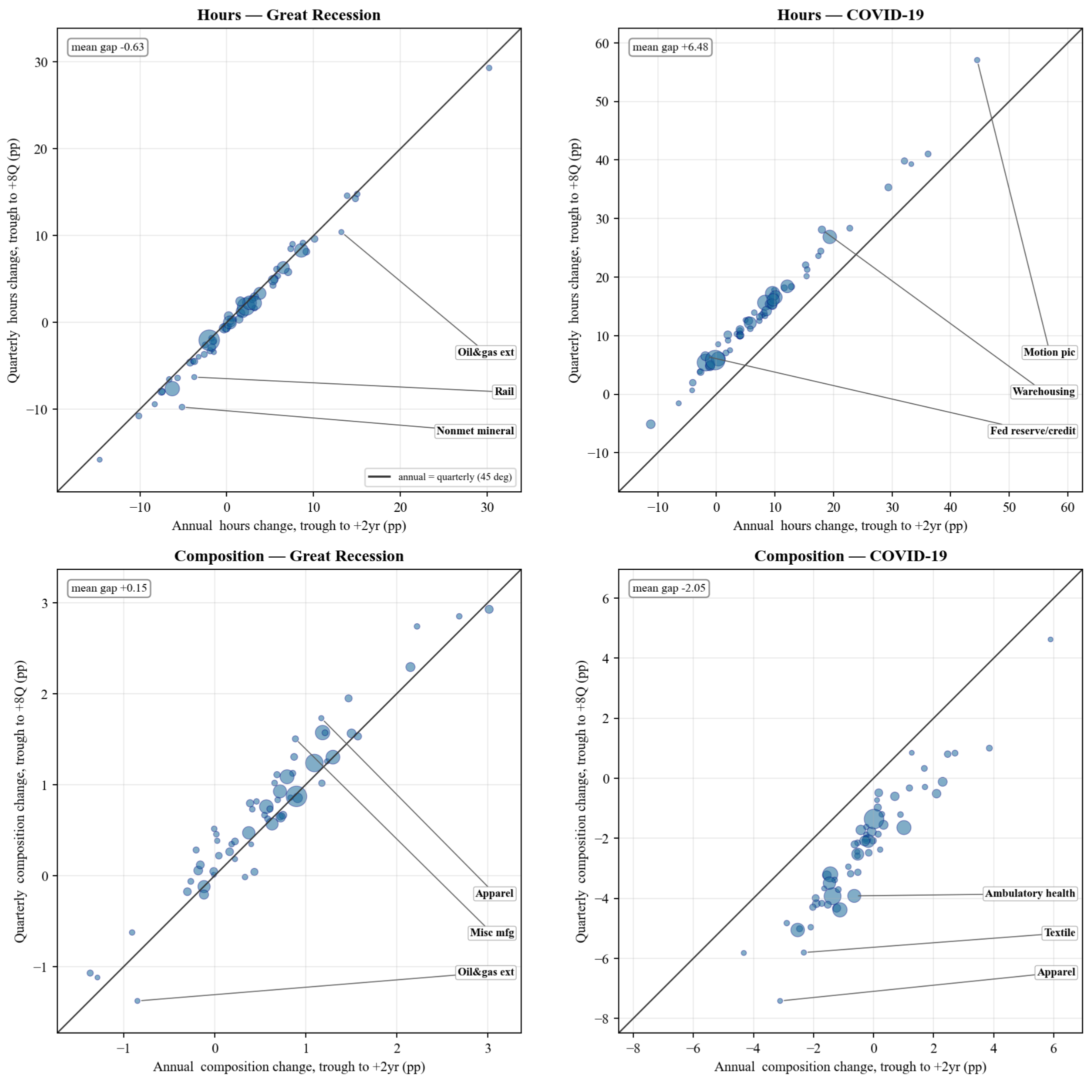

Figure 6. Recovery changes over the two years after the trough, quarterly (vertical) versus annual (horizontal), for raw hours (top row) and within-industry composition (bottom row), in the Great Recession and COVID-19. Each point is one of the 63 industries (finance industries 42 and 44 dropped), and marker size is the industry's hours share. The 45-degree line marks equality of the two frequencies, and each panel reports the mean quarterly-minus-annual gap. In COVID-19 the hours points lie above the line, as annual data understate the rebound, and the composition points lie below it, as annual data miss the within-year reversal of the spike. In the Great Recession both margins lie close to the line.

2021-2022 slowdown, likely reflect who was working rather than changes in efficiency. Annual labor-quality estimates, which understate the spike at its peak and cannot see it reverse, allocate this margin to different years than the quarterly series would. A quarterly quality-adjusted account would therefore discipline interpretations of the post-pandemic productivity path that currently rest on annual composition adjustment.

Real-time tracking of labor composition would also inform policy during rapid shocks. Aggregate averages such as hourly earnings and output per hour moved sharply in 2020 partly because the composition of the employed workforce changed, rather than solely because individual wages or productivity changed. A quarterly composition measure separates these margins as events unfold. For monetary and fiscal authorities assessing labor-market slack, wage pressure, or potential output in the middle of a fast-moving downturn, knowing whether a surge in average wages or measured productivity is a composition artifact of who lost employment is directly decision-relevant. The present paper's central contribution is the quarterly measurement of labor input and its cyclical composition.

A. CPS Variable Definitions and Worker-Type Classifications

Table A.1 documents the worker-type groupings constructed from CPS microdata, and Table A.2 documents the CPS source variables and construction rules. The four additive labor input quantities, employment (E), employment-hours (EH), employment-weeks (EW), and total labor compensation ($EHWC$), are aggregated to each cell defined by the cross-product of these dimensions interacted with 63 BEA industry groups.

B. Detailed Methodology

B.1. 63 Industry to 13 Sector Mapping

Table B.1 lists the mapping from the 63 BEA industries to the 13 broad sectors used in the marginal construction.

Table A.1. **Worker-Type Cell Boundaries**

Dimension	Group	Label	CPS / IPUMS criterion
Age (`agegrp`)	1	1-15	age ≤ 15; zeroed out before labor input aggregation
	2	16-17	$16 \leq$ age ≤ 17
	3	18-24	$18 \leq$ age ≤ 24
	4	25-34	$25 \leq$ age ≤ 34
	5	35-44	$35 \leq$ age ≤ 44
	6	45-54	$45 \leq$ age ≤ 54
	7	55-64	$55 \leq$ age ≤ 64
	8	65+	age ≥ 65
Education (`edugrp`)	1	Less than 9th grade	educ ≤ 39
	2	9th-12th grade, no diploma	$40 \leq$ educ ≤ 72
	3	High school diploma / GED	educ $= 73$
	4	Some college or associate degree	$80 \leq$ educ ≤ 122, excluding 111
	5	Bachelor's degree	educ $= 111$
	6	Graduate degree (master's, prof., doctoral)	educ ≥ 123
Sex (`sex`)	1	Male	`sex` $= 1$
	2	Female	`sex` $= 2$
Worker class (`classgrp`)	1	Wage and salary (private and government)	$15 \leq$ `classwkr` ≤ 25 or $27 \leq$ `classwkr` ≤ 28
	2	Self-employed	$10 \leq$ `classwkr` ≤ 14

Table A.2. CPS Source Variables and Construction of the Additive Quantities

Variable	CPS source	Construction
Employment (E)	wtfinl$/3$	Weighted count of employed persons (empstat 10, 12). The weight is divided by 3 to convert monthly to quarterly
Hours (EH)	uhrswork1	$\sum$ uhrswork1 $\times$ (wtfinl$/3$). Usual hours worked per week, with values ≥ 900 set missing
Weeks (EW)	wksworkorg	Weeks worked per year, deliberately harmonized to interval midpoints in the style of the Annual Social and Economic Supplement (ASEC) so the quarterly indicator matches the annual benchmark's weeks definition. Cell-level W is the weighted mean among valid reporters, and the benchmarking quantity EW is constructed as $E \times W$, so reporting rates do not enter the indicator's variance. Reported for a subset of ORG earners (see notes)
Weekly earnings (EHC)	earnweek	$\sum$ earnweek $\times$ (earnwt$/3$). Usual weekly earnings, with topcoded values replaced with the mean above the topcode implied by a fitted log-normal earnings distribution. Hourly compensation C is derived as weekly earnings divided by average weekly hours. The benchmarking quantity $EHWC$ is then constructed as the product $E \times H \times W \times C$, not by summing earnweek directly

Notes: Industry assignment follows four crosswalk vintages (2003-2008, 2009-2013, 2014-2019, and 2020 onward), reflecting NAICS reclassification cycles. Federal government workers (classwkr $= 25$) are assigned to industry 62, and state and local government workers (classwkr $\in \{27, 28\}$) to industry 63, overriding the crosswalk assignment. The topcode threshold is \$2,884.61/week. Topcoded observations are replaced with the conditional mean above the topcode estimated under a log-normal model of weekly earnings, following the Center for Economic and Policy Research CPS-ORG methodology (Schmitt, 2003). From 2023Q2 onward, IPUMS sources weekly earnings from the redesigned EARNWEEK2 variable, which carries month-varying swap-value topcodes and a different not-in-universe code. These observations are harmonized to the legacy coding (not-in-universe recoded, values above \$2,884.61 re-capped) so that the fixed-topcode log-normal correction applies uniformly across 2003-2024. Valid weeks-worked reports cover about 26 percent of the ORG earner universe (5.4 percent of all employed respondents over 2003-2024). Cell-level W is computed as the mean among reporters, and because its level and trend are anchored to annual ILPA weeks through Denton-Cholette benchmarking, with the quarterly indicator heavily smoothed, the restricted reporting universe affects only the within-year shape of the weeks path.

Table B.1. Broad sectors used in aggregation

Code	Sector	Industry codes (1-63)
1	agriculture	1-2
2	mining & utilities	3-6
3	construction	7
4	durable manufacturing	8-18
5	nondurable manufacturing	19-26
6	trade	27-28
7	transportation	29-36
8	information & finance	37-45
9	professional & business services	46-52
10	education & health	53-56
11	leisure & hospitality	57-60
12	other services	61
13	government	62-63

Note: These 13 broad sectors aggregate the 63 BEA detailed industries using the code ranges shown. Industry definitions follow the standard BEA industry classification.

B.2. Marginal Creation

The first stage constructs annual marginal benchmarks from the published ILPA data. Marginals are dimensional aggregates computed across each of the five marginal vectors, serving as targets for subsequent temporal disaggregation and distribution adjustment.

We construct the four additive quantities employment E, employment-hours EH, employment-weeks EW, and total labor compensation $EHWC$ across five mutually consistent marginal vectors:

- Worker types (education, age, sex), $6 \times 8 \times 2 = 96$ cells
- Industry, 63 NAICS groups
- Class and sector, $2 \times 13 = 26$ cells
- Education and sector, $6 \times 13 = 78$ cells

- Sex and sector, $2 \times 13 = 26$ cells

Because the Denton-Cholette method operates on additive quantities, we compute multiplicative aggregates (EH, EW, EHWC) to preserve the accounting identity $E \times H \times W \times C =$ Total Compensation.

Quarterly marginal aggregates are then seasonally adjusted using X-13ARIMA-SEATS (U.S. Census Bureau, 2017) to remove seasonal patterns, producing seasonally adjusted series that isolate trend and cyclical variation. The adjustment uses an additive X-11 decomposition with no transformation, and the regARIMA stage's automatic outlier detection remains enabled, so extreme pandemic observations are captured by automatically identified additive-outlier and level-shift regressors rather than absorbed into the seasonal factors. This prevents the 2020Q2 collapse from distorting the estimated seasonal pattern. For the small number of series where X-13 estimation fails, the pipeline falls back to the unadjusted series and logs the case. Seasonal adjustment precedes state space filtering to prevent the Kalman smoother from attributing seasonal fluctuations to trend volatility.

B.3. Quarterly Smoothing via State Space Modeling with Annual Dynamics

The smoothing procedure has two steps. It estimates trend variances from the cleaner annual ILPA benchmarks and applies them to smooth the quarterly indicators.

B.3.1. Learning Dynamics from Annual Data

For each annual ILPA marginal series, we fit a local linear trend state space model in which the log of the series is a noisy observation of a level that follows a random walk with stochastic drift:

$$\ln(X_y^{\text{annual}}) = \mu_y + \epsilon_y, \qquad \epsilon_y \sim N(0, R), \tag{B.1}$$

$$\mu_{y+1} = \mu_y + \beta_y + \eta_{\mu,y}, \qquad \eta_{\mu,y} \sim N(0, Q_\mu), \tag{B.2}$$

$$\beta_{y+1} = \beta_y + \eta_{\beta,y}, \qquad \eta_{\beta,y} \sim N(0, Q_\beta), \tag{B.3}$$

where μ_y is the level and β_y is the slope. Annual equations are indexed by y and quarterly equations below by q.

We estimate the process variances Q_μ and Q_β (representing the true trend volatility in annual data) and measurement-error variance R by maximum likelihood using the Kalman filter, implemented with the KFAS state-space package (Helske, 2017) in R (R Core Team, 2024).[3]

B.3.2. Applying Dynamics to Quarterly Smoothing

The estimated quarterly state space model uses the trend dynamics learned from annual data:

$$\ln(X_q^{\text{quarterly}}) = \mu_q + \epsilon_q, \tag{B.4}$$

with process variances inherited directly from the annual estimates:

$$Q_\mu^{\text{quarterly}} = Q_\mu^{\text{annual}}, \quad Q_\beta^{\text{quarterly}} = Q_\beta^{\text{annual}}. \tag{B.5}$$

This transfer is a deliberate heuristic, not a temporal-aggregation identity. Because innovation variances cumulate over time, a frequency-consistent local-linear-trend mapping would scale the quarterly variances below the annual ones (for the level, by roughly the number of quarters per year). Carrying the annual variances over unscaled makes the quarterly trend more responsive than strict aggregation arithmetic would imply. We adopt the unscaled transfer for two reasons. First, the filter's behavior is governed by the signal-to-noise ratio Q/R rather than by Q alone, and the quarterly measurement variance $R^{\text{quarterly}}$ is estimated separately (with the safeguards described below), so the effective degree of smoothing is disciplined on the noise side. Second, any residual over-responsiveness in the smoothed indicator affects only within-year shape. The subsequent Denton-Cholette stage pins every year's level to the annual ILPA benchmark, so the choice does not affect annual levels. In practice the unscaled transfer errs toward preserving within-year movement, which is the object of interest, and the benchmarking stage bounds the cost of that

[3] Unlike some implementations that fix $R = 0$ for annual data, we estimate R from the annual series. This allows the model to distinguish between genuine trend variation (captured by Q) and observational noise (captured by R), preventing inflated Q estimates from conflating true variation with measurement error.

choice.

The quarterly measurement-error variance $R^{\text{quarterly}}$ is estimated separately from the seasonally adjusted quarterly CPS microdata, reflecting the higher measurement error in quarterly survey data versus annual benchmarks.[4]

We apply the Kalman smoother (KFS) to obtain smoothed state estimates for the quarterly series. The smoothed level component μ_q represents the underlying quarterly trend with noise removed, guided by the dynamics learned from the more reliable annual data.

*Exception: worker-type employment.*Employment for the worker-type marginal (education $\times$ age $\times$ sex) bypasses the Kalman step. Its seasonally adjusted series is used directly as the Denton-Cholette indicator. As stated in Section 4, the purpose of the exception is to keep aggregate labor-composition change timely. The worker-type cells are the direct source of aggregate labor-composition movement, so smoothing this indicator damps and delays the aggregate composition path, including the 2020Q2 reallocation that is central to the results. Bypassing the filter is acceptable for this series because employment is collected for every CPS respondent, unlike weeks and earnings, which come only from the outgoing rotation groups, and the worker-type marginal pools the full sample across all industries, so its quarterly movements are the most reliably measured in the system. Consistent with this, the unsmoothed treatment keeps aggregate labor-composition growth in line with external estimates constructed directly from CPS demographic shares, which apply no such filter. Hours, weeks, and compensation on every marginal, and employment on the other marginal vectors, retain the full state-space treatment described above. The unsmoothed employment indicator is still benchmarked to annual ILPA means in Stage 1, so annual consistency is unaffected.

B.3.3. Sequential Two-Stage Calibration

The quarterly labor input estimates are subjected to sequential two-stage calibration, first to annual ILPA benchmarks through Denton-Cholette temporal disaggregation and then to quarterly CPS aggregate employment through proportional scaling. The two stages serve different purposes. Denton-Cholette anchors the smoothed quarterly path to the annual

[4]As a sanity check that quarterly survey data are not treated as more precise than annual data, we require $R^{\text{quarterly}} \geq R^{\text{annual}}$. When the freely estimated quarterly variance falls below the annual estimate, it is reset to $2 \times R^{\text{annual}}$. Additionally, if the signal-to-noise ratio $Q_\mu^{\text{quarterly}}/R^{\text{quarterly}}$ exceeds 0.05, we adjust $R^{\text{quarterly}}$ upward to prevent the filter from tracking noise rather than smoothing.

ILPA account, while the CPS proportional adjustment aligns the marginal-stage quarterly employment level with the high-frequency aggregate employment series. Because the CPS adjustment is applied after annual benchmarking rather than solved jointly with it, exact annual aggregation holds after Stage 1 and the Stage 2 employment series should be interpreted as CPS-aligned quarterly estimates. A third, cell-level alignment applied after IPF (Appendix B.5) subsequently restores annual ILPA consistency in the final detailed data. This approach balances two objectives:

1. **Annual benchmarking to ILPA**: The official ILPA data provides authoritative annual estimates of employment, hours, weeks, and compensation by industry and worker-type group. The Stage 1 quarterly estimates are benchmarked to these annual means before the CPS proportional adjustment.

2. **Quarterly alignment with CPS**: The CPS is the timeliest source of high-frequency employment movement. Each marginal vector's quarterly employment total is aligned to the pipeline's X-13 seasonally adjusted aggregate of CPS employment, computed within that vector, so that the quarterly path of employment reflects the high-frequency movement of the survey.

The first stage imposes annual ILPA consistency, and the second stage aligns quarterly employment with CPS. This approach preserves the relative distribution of employment across worker-type and industry groups (important for the subsequent IPF step) while anchoring the quarterly path of employment to the CPS-based seasonally adjusted aggregate.

B.3.4. Stage 1: Denton-Cholette Benchmarking

Following the smoothing stage, let I_{yq} denote the quarterly indicator for quarter q of year y, meaning the Kalman-smoothed series in general and the seasonally adjusted series for worker-type employment, and let x_{yq} denote the benchmarked quarterly estimate. We apply Denton-Cholette temporal disaggregation using I_{yq} as the high-frequency indicator and the corresponding annual ILPA benchmark X_y^{ILPA} as the low-frequency constraint. Operationally, this is implemented using the `tempdisagg` package (Sax and Steiner, 2013) with `method = "denton-cholette"` and `conversion = "mean"`, so that quarterly estimates preserve the movement in the Kalman-smoothed indicator while remaining consistent with annual benchmark values. The binding low-frequency constraint is therefore

$$\frac{1}{4}\sum_{q=1}^{4} x_{yq} = X_y^{\mathsf{ILPA}}, \tag{B.6}$$

not a four-quarter sum constraint.

In the main product-based benchmarking path, Denton-Cholette is applied to additive quantities rather than only to component averages. Specifically, we benchmark E, EH, EW, and $EHWC$ to their corresponding annual ILPA marginal means, where $EH = \sum EH$, $EW = \sum EW$, and $EHWC = \sum EHWC$ are constructed as sum-of-products at the underlying cell level. On the quarterly side, the high-frequency indicators for these products are constructed by multiplying the corresponding component indicators together prior to benchmarking, $I_{EH} = I_E \times I_H$ and analogously for EW and $EHWC$. Here I_E, I_H, I_W, and I_C are the component indicator series defined above, one per variable with the time subscripts suppressed. I_E is Kalman-smoothed except for the worker-type marginal, where the seasonally adjusted series enters directly, and I_H, I_W, and I_C are Kalman-smoothed throughout. The products themselves are never run through the Kalman filter. This construction keeps the component recovery below well behaved. Because the same employment indicator enters both the benchmarked product and benchmarked employment, dividing one by the other cancels the common employment factor, and the recovered rates inherit the smoothness of the rate indicators. The component variables are then recovered as:

$$H_q = \frac{(EH)_q}{E_q}, \qquad W_q = \frac{(EW)_q}{E_q}, \qquad C_q = \frac{(EHWC)_q}{E_q H_q W_q}. \tag{B.7}$$

This product-based benchmarking preserves the accounting identities by construction, while the Kalman-smoothed quarterly series governs the within-year movement of the benchmarked path.

When any step of this path fails for a series, the pipeline falls back conservatively, benchmarking the available component variables individually or retaining the prior stage's series, and logs the case for diagnostic review.

B.3.5. Stage 2: Quarterly Proportional CPS Adjustment

Following Stage 1 Denton-Cholette benchmarking, Stage 2 applies a second calibration step that aligns quarterly employment marginals to aggregate CPS data. This step ensures that the sum of employment across each marginal vector matches that vector's seasonally adjusted CPS employment aggregate, constructed within the pipeline by X-13 adjustment of the vector's own CPS tabulations.

Let $E^{\mathsf{CPS}}_{s,q}$ denote the seasonally adjusted CPS employment aggregate for marginal vector s in quarter q, computed within the pipeline as the sum of the vector's X-13-adjusted cell indicators, and let $\tilde{E}_{m,q}$ denote the Denton-benchmarked employment of cell m within marginal vector s. Because Stage 2 precedes the IPF module, the detailed worker-type-by-industry array does not yet exist. The scaling operates on each marginal vector separately. A proportional scaling factor is computed as:

$$\lambda_{s,q} = \frac{E^{\mathsf{CPS}}_{s,q}}{\sum_{m\in s} \tilde{E}_{m,q}}. \tag{B.8}$$

Employment for each marginal cell is then scaled by this factor:

$$E_{m,q} = \lambda_{s,q} \cdot \tilde{E}_{m,q}. \tag{B.9}$$

This adjustment ensures that quarterly employment marginals reflect CPS aggregate employment levels while preserving the relative distribution of employment across worker-type and industry groups within each quarter. The tradeoff is that, because scaling is applied after Denton-Cholette benchmarking, the adjusted quarterly employment series does not aggregate exactly back to the preadjustment annual ILPA employment means. Quantitatively, the wedge is computed for each year of 2003-2024 as the percentage deviation of the annual mean of the CPS-aligned quarterly aggregate employment from the corresponding ILPA total employment. It averages -5.3 percent, ranging from -4.3 to -6.4 percent. The year-to-year change in the wedge averages 0.3 percentage point, so the within-year shape carried forward is essentially unaffected. The wedge is confined to this intermediate stage. The annual cell-level alignment applied after IPF (Appendix B.5) restores ILPA annual consistency in the final detailed data. The scaling is applied only to employment (E). Hours (H), weeks (W), and compensation (C) retain their Denton-benchmarked values from the prior step, so the multiplicative identities are maintained using the scaled employment component and the unchanged hours, weeks, and compensation components.

B.4. Distribution Matching via IPF

After quarterly marginal series are established through seasonal adjustment, Kalman smoothing, and Denton-Cholette benchmarking, the annual ILPA cell distribution is adjusted to the quarterly marginals using a sequence of IPF procedures.

*Marginal harmonization.*Two preparatory steps make the five vectors' targets mutually consistent before any IPF pass. First, each quarter's employment totals for the industry, class-by-sector, education-by-sector, and sex-by-sector vectors are iteratively rescaled to equal the worker-type vector's total (tolerance 10^{-5}). One-shot analogues align the EH, EW, and $EHWC$ marginals to totals derived from the same worker-type reference. Second, for every product, the education-by-sector and sex-by-sector targets are reconciled by two-dimensional IPF (tolerance 10^{-4}, up to 500 iterations) to education and sex controls computed from the worker-type vector and to sector controls computed from the industry vector via the industry-to-sector mapping, and are then rescaled to the industry-derived grand total. The class-by-sector targets are scaled sector-by-sector to the same controls. The worker-type and industry vectors thus serve as the level authorities for the demographic and sector dimensions, respectively. Without this harmonization the five separately smoothed and benchmarked vectors would present the IPF loop with mutually inconsistent targets. The reconciliation itself then proceeds product by product, as described below.

B.4.1. Two-Dimensional RAS as an IPF Special Case

In two dimensions, the RAS algorithm, a special case of IPF, adjusts a matrix $\mathbf{X}^{\text{seed}}$ to satisfy specified row and column sums through sequential (alternating) proportional scaling. Within each iteration, rows are scaled first, and columns are then scaled using the already row-updated values:

$$X_{ij}^{(n+\frac{1}{2})} = X_{ij}^{(n)} \cdot \frac{r_i}{\sum_{j'} X_{ij'}^{(n)}} \tag{B.10}$$

$$X_{ij}^{(n+1)} = X_{ij}^{(n+\frac{1}{2})} \cdot \frac{c_j}{\sum_{i'} X_{i'j}^{(n+\frac{1}{2})}} \tag{B.11}$$

The algorithm converges when row and column totals satisfy:

$$\sum_j X_{ij} = r_i \quad \forall i \tag{B.12}$$

$$\sum_i X_{ij} = c_j \quad \forall j \tag{B.13}$$

within a specified tolerance.

B.4.2. Multivariate Application

The IPF adjustment operates over a five-dimensional structural array indexed by industry (63 groups), sex (2), age (8), education (6), and class-of-worker (2). The youngest age group, ages 1-15, is retained in this structural array but is zeroed out before labor input aggregation, leaving 168 active worker types. Within each iteration, five sequential scaling passes are applied in the following order:

1. Scale to match demographic marginals (education × age × sex)
2. Scale to match industry marginals
3. Scale to match class-of-worker × sector marginals
4. Scale to match education × sector marginals
5. Scale to match sex × sector marginals

Each pass applies the sequential proportional sweep described in the preceding subsection across the relevant dimensions, leaving all other dimensions unchanged. For example, the industry scaling in step 2 computes, for each industry i:

$$X_{g,i}^{(*)} = X_{g,i} \cdot \frac{E_i^{\text{target}}}{\sum_g X_{g,i}}, \tag{B.14}$$

where E_i^{target} is the quarterly industry marginal from the state space smoothing step. For E and EH, all five passes are repeated until the applicable constraints are satisfied within a tolerance of 10^{-4}. For EW and $EHWC$, only the worker-type, education-by-sector, and sex-by-sector passes are binding inside the iterative loop. Industry and class-by-sector controls are pre-reconciled but not repeatedly imposed.

B.4.3. Multiplicative Decomposition for H, W, and C

IPF is applied sequentially to employment and to composite quantities. The first pass reconciles employment (E). The second pass reconciles employment-hours (EH), constructed using the IPF-adjusted employment base and the quarterly hours information from the smoothed marginal series. The third pass reconciles employment-weeks (EW) analogously. The fourth pass reconciles total labor compensation ($EHWC$), using the IPF-adjusted employment, implied hours, implied weeks, and hourly compensation information. After each composite IPF pass, the corresponding per-worker quantity is recovered as an implied ratio:

$$H_{g,i} = \frac{(EH)_{g,i}}{E_{g,i}} \tag{B.15}$$

$$W_{g,i} = \frac{(EW)_{g,i}}{E_{g,i}} \tag{B.16}$$

$$C_{g,i} = \frac{(EHWC)_{g,i}}{E_{g,i} H_{g,i} W_{g,i}}. \tag{B.17}$$

This sequence preserves the $E \times H \times W \times C$ accounting identity by construction. Employment is first reconciled directly, hours and weeks are recovered from IPF-balanced product totals, and compensation is recovered from the reconciled total labor compensation aggregate relative to the implied EHW base.

For the EW and $EHWC$ passes, the iterative loop actively enforces only constraints 1 (education $\times$ age $\times$ sex), 4 (education $\times$ sector), and 5 (sex $\times$ sector). This reflects the lower cell-level sample counts for weeks and compensation in the CPS microdata at fine industry and class-of-worker breakdowns. The IPF adjustment reconciles the detailed distribution to the applicable quarterly targets while preserving the worker-type structure of the annual ILPA seed.

B.5. Annual Cell-Level Alignment

After IPF, a final calibration aligns the detailed cells to the annual ILPA account. For each variable, the four quarterly IPF results for a year are loaded and, for every detailed worker-type-by-industry cell, the quarterly values are averaged, matching the annual-mean convention of the Denton stage. A single scale factor per cell and year is computed as the ratio of the ILPA cell benchmark to this quarterly average and applied to all four quarters, so within-year movement is preserved exactly while the cell's annual mean is restored to the benchmark. Edge cases are handled conservatively. Cells with missing data, or with zero in both the quarterly average and the benchmark, are left unchanged. Cells with a zero quarterly average but a positive benchmark are left at zero, since all four quarters are zero in that case. The proportional factor is undefined there, and rather than invent a within-year shape, the alignment leaves those cell-year benchmarks unmet. Cells with positive quarterly values but a zero benchmark are forced to zero. Employment is aligned to the ILPA employment cell benchmark and validated to reproduce it to within 10^{-3}. Hours and weeks are aligned to employment-weighted benchmark analogues, and hourly compensation to an employment-hours-weeks-weighted analogue, consistent with the aggregation conventions of Section 3. All labor-composition results in the paper are computed from the aligned cells.

B.6. Implementation Notes

This subsection documents implementation choices that the algorithm makes for numerical stability and data quality, beyond the core methodology described above.

B.6.1. Zero Handling for Cell-Level Series

Quarterly CPS marginal series contain frequent zeros from sparse marginal cells (for example, thin education-by-sector crosses), and these zeros require different treatment depending on the variable. For employment (E), all zeros are replaced by 10^{-9} before log transformation, treating them as sparse but nonmissing cells. For hours (H), weeks (W), and compensation (C), runs of two or more consecutive zeros are treated as genuine missing periods and set to NA, while isolated single zeros are treated as sparse observations and set to 10^{-9}. This differentiation avoids the alternative of either (a) converting all zeros to 10^{-9}, which creates implausibly large negative log values for W, H, and C, or (b) converting all zeros to NA, which discards sparse-but-valid observations and inflates state-space variance estimates.

B.6.2. Class-2 Imputation for W and C in the Class-by-Sector Marginal

The CPS earner study does not collect weekly earnings or weeks worked for the self-employed, so weeks worked per year (W) and hourly compensation (C) are unavailable for the self-employed class. Government employees are part of the wage-and-salary class, and their reported earnings are used directly. For the class-by-sector marginal, we impute the self-employed class's values of W and C from the wage-and-salary class within the same sector and quarter before applying the seasonal adjustment, Kalman smoothing, and Denton-Cholette pipeline. This ensures the pipeline operates on realistic non-zero values rather than artifacts of CPS coverage.

A second, cell-level imputation occurs in the labor-composition calculation itself. After the annual alignment, each self-employed cell's hourly compensation and weeks are replaced with the wage-and-salary values from the same industry $\times$ age $\times$ education $\times$ sex cell before the Törnqvist compensation shares are computed. Because the CPS collects no earnings information for the self-employed, the pipeline-implied values for these cells derive entirely from the marginal-level imputation above. The cell-level replacement makes the treatment explicit and uniform at the level where cost shares are formed.

B.6.3. Physical-Range Clamping

Pipeline outputs are clamped to physically valid ranges to prevent occasional numerical divergence of the Kalman smoother on sparse cells from corrupting downstream estimates. The ranges are weeks per year $W \in [0, 52]$, hours per week $H \in [0, 168]$, employment

$E \geq 0$, and compensation $C \geq 0$.

B.6.4. Annual Employment Interpolation for the IPF Seed

The IPF algorithm requires a seed matrix describing the worker-type-by-industry cell distribution. We use the annual ILPA cell counts as the seed but linearly interpolate adjacent annual values to obtain a quarterly-appropriate seed. Treating each annual value as centered on the year's midpoint, Q1 receives 0.375 weight from the previous year and 0.625 from the current year, Q2 receives 0.125 from the previous year and 0.875 from the current year, and Q3 and Q4 mirror these weights with the next year. This interpolation prevents discrete year-boundary jumps in the IPF seed that would otherwise propagate to the quarterly cell distributions.

References

Mark J. Bils. Real wages over the business cycle: Evidence from panel data. *Journal of Political Economy*, 93(4):666–689, 1985.

Gregory C. Chow and An-loh Lin. Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *Review of Economics and Statistics*, 53 (4):372–375, 1971.

Estela Bee Dagum and Pierre A. Cholette. *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*, volume 186 of *Lecture Notes in Statistics*. Springer, New York, 2006.

Frank T. Denton. Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, 66 (333):99–102, 1971.

John G. Fernald. A quarterly, utilization-adjusted series on total factor productivity. Working Paper 2012-19, Federal Reserve Bank of San Francisco, 2014.

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Etienne Breton, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, David Van Riper, and Kari C.W. Williams. IPUMS CPS: Version 13.0 [dataset]. University of Minnesota: IPUMS, Minneapolis, MN, 2025. URL https://doi.org/10.18128/D030.V13.0.

Corby Garner, Justin Harper, Thomas F. Howells, Matthew Russell, and Jon D. Samuels. Integrated Industry-Level Production Accounts for the United States: Experimental Statistics for 1987-1997, Revised Statistics for 1998-2015, and Initial Statistics for 2016. *Survey of Current Business*, 98(7):1–28, July 2018.

Jouni Helske. KFAS: Exponential family state space models in R. *Journal of Statistical Software*, 78(10):1–39, 2017.

Bart Hobijn, Martí Mestieri, Nicolas Werquin, and Jing Zhang. Quarterly industry-level labor productivity data for the U.S. *Federal Reserve Bank of Chicago, Economic Perspectives*, (1), 2025.

Dale W. Jorgenson and Zvi Griliches. The explanation of productivity change. *Review of Economic Studies*, 34(3):249–283, 1967.

Dale W. Jorgenson, Frank M. Gollop, and Barbara M. Fraumeni. *Productivity and U.S. Economic Growth*. Harvard University Press, Cambridge, MA, 1987.

Dale W. Jorgenson, Mun S. Ho, and Kevin J. Stiroh. *Productivity, Volume 3: Information Technology and the American Growth Resurgence*. MIT Press, Cambridge, MA, 2005.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2024.

Christoph Sax and Peter Steiner. Temporal disaggregation of time series. *The R Journal*, 5(2):80–87, 2013.

John Schmitt. Creating a consistent hourly wage series from the Current Population Survey's Outgoing Rotation Group, 1979–2002. Working paper, Center for Economic and Policy Research, Washington, DC, 2003.

Gary Solon, Robert Barsky, and Jonathan A. Parker. Measuring the cyclicality of real wages: How important is composition bias? *Quarterly Journal of Economics*, 109(1): 1–25, 1994.

Jay Stewart. Why was US labor productivity growth so high during the COVID-19 pandemic? The role of labor composition. *International Productivity Monitor*, 42:87–103, 2022.

U.S. Census Bureau. *X-13ARIMA-SEATS Reference Manual, Version 1.1*. U.S. Census Bureau, Time Series Research Staff, Washington, DC, 2017.