METHODOLOGICAL APPENDIX

Core labor market information

Brookings Metro's primary provider of core labor market information is Lightcast (formerly known as Emsi Burning Glass). This analysis uses Lightcast data for all information pertaining to employment, wages, output, and population, with disaggregation by age, educational attainment, sex, race, detailed industry, and detailed occupation where available. For more detail on Lightcast's data collection and modeling procedures, visit https://kb.lightcast.io.

Explaining shift-share

This market assessment uses shift-share analysis to identify industrial competitive advantages over time, holding national growth rates constant. Shift-share (competitive effect) analyses are conducted as a measure of the competitiveness of an industry's growth trajectory in a given region, and are not indicative of whether an industry is growing or shrinking. An industry that has gained jobs, output, or payroll over a period but demonstrated a negative competitive effect is an industry in which inherent characteristics or market conditions within a given region are causing it to underperform the nation. Conversely, an industry that appears to be shrinking but demonstrates a positive competitive effect is an industry in which local/regional conditions are mitigating the impacts of national decline.

The competitive effect for each industry is defined as the difference in the actual and expected change in an indicator over time, where expected change is the sum of the industry mix effect (each industry's national growth rate less the national growth rate for all industries, multiplied by the base-year value) and the national growth effect (the national growth rate for all industries, multiplied by the base-year value). Alternatively, for summative indicators including but not limited to employment, output, and total payroll:

$$ce_{i,j(y-y_b)} = \left(x_{i,j,y} - x_{i,j,y_b}\right) - \left(x_{i,j,y_b}\left(\frac{x_{i,u,y} - x_{i,u,y_b}}{x_{i,u,y_b}} - \frac{x_{t,u,y} - x_{t,u,y_b}}{x_{t,u,y_b}}\right) + x_{i,j,y_b}\left(\frac{x_{t,u,y} - x_{t,u,y_b}}{x_{t,u,y_b}}\right)\right),$$

Where x represents the value of a given indicator, i represents each industry, t represents the total for all industries, j represents the region of interest, u represents the nation, y represents the analysis year, and y_b represents the base year. This report uses 2023 as the analysis year and 2012 as the base year, except where otherwise specified. For non-summative/ratio indicators (such as average earnings and productivity), each counterfactual decomposition is modeled from the decomposition values of each ratio input, weighted by employment.

In traditional shift-share analyses, expected growth and competitive effect levels are calculated independently at different levels of industry aggregation. While this is often the preferred approach for public datasets in which data in some industries may be incomplete or suppressed, this approach ignores intra-sector variation in industry growth rates, leading to inconsistent competitive effect values. To preserve the decomposition logic of shift-share, we apply the above equation to data at the most granular industry level (NAICS-6) and aggregate upward to determine the competitiveness of summative indicators (e.g., employment, output, and total payroll) for higher-level sectors and sub-sectors. This bottom-up approach to shift-share avoids masking variation in industries where regions have a very high or very low degree of specialization relative to the nation.

Classes of workers

All employment estimates are based on Lightcast's full labor market dataset and include four classes of workers:

- ***QCEW-eligible employment:*** Jobs eligible for state and local unemployment insurance (UI) programs, accounting for more than 95% of employees who receive regular compensation from an employer and are subject to federal, state, and local taxes. Excludes most agricultural workers, railroad employees, self-employed workers, and workers at some tax-exempt nonprofits.

- ***Non-QCEW employment:*** Wage and salary (employer-based) jobs that are not eligible for unemployment insurance, representing approximately 5% of the U.S. workforce. Includes certain agricultural workers, railroad employees, state and local government workers, domestic workers, and workers of certain national security agencies.

- ***Self-employed workers:*** Includes workers who earn income primarily from self-employment, such as through business ownership.

- ***Extended proprietors:*** Workers who report miscellaneous earnings outside of their primary source of income, including (but not limited to) creative workers and workers in the gig economy.

There are limitations in including extended proprietors in employment data, since these workers are identified based on supplemental income reported to the IRS (and definitionally receive the majority of their income from other sources, including traditional wage employment). We include all classes of workers, including extended proprietors, for the following reasons:

- It is not possible to fully differentiate between QCEW-eligible and non-QCEW workers, and between self-employed workers and self-proprietors, in person-level datasets such as the American Community Survey and Current Population Survey where individuals denote more than one labor income source outside of wage and salary employment. The inclusion of all workers allows us to more effectively weight workers based on their occupational classifications in these datasets. Similarly, the resume data undergirding our analysis of occupational transitions and career mobility cannot account for different classes of workers, making it impossible to accurately assess the differences in career mobility between workers of different classes in the same occupation.

- Extended proprietor earnings are critical in the input-output models used by Lightcast and other economic modeling agencies (such as RIMS-II and IMPLAN) to calculate gross domestic product. As a result, calculations of worker productivity and regional standards of living that do not account for extended proprietors may be inflated.

- Percentile wage distributions at and below the median are nearly always higher for QCEW-eligible workers than for extended proprietors. Because of this, restricting the dataset to exclude self-employed workers and extended proprietors may overstate wage expectations for certain occupations. Including self-employed and extended proprietors provides a more conservative approach to estimating job quality and career mobility.

When interpreting data including extended proprietors, it is important to note that estimates are representative of the total number of jobs, not the total number of workers, and that employment may be overrepresented in some sectors. Where applicable, this market assessment provides separate estimates for QCEW-eligible workers to enable benchmarking against federal data sources and provide a more accurate assessment of firm-based industry characteristics. While the QCEW accounts for more than 95% of firm-based employment, these estimates may undercount agricultural and farm workers, railroad employees, and employees at tax-exempt nonprofit and religious institutions.

Struggling families, living wage standards, and opportunity jobs

One primary objective of this market assessment is to identify opportunities for industry growth in Southeastern Pennsylvania that support economic self-sufficiency and upward mobility for the region's workers. A key component of such analyses is the identification of a living wage standard, which—though ostensibly a core component of most frameworks for "good jobs"—is often directly anchored to family sizes, and thus does not translate well into demand-side strategies for employment growth, talent attraction, and workforce development.

To mitigate this limitation, Brookings Metro uses family-based cost-of-living thresholds provided by the University of Washington's Center for Women's Welfare and the Economic Policy Institute's Family Budget Calculator to create a composite cost-of-living threshold that would provide wage sufficiency for the majority of families in each U.S. county without being defined by a single family archetype. Each composite threshold is based on the median gross total cost of living for housing, child care, food, transportation, health care, emergency savings, miscellaneous necessities, and taxes. We use the average ratio of emergency savings to the total cost of living (less taxes) as a proxy in counties where data on emergency savings is unavailable. For all counties, we add 10% of the pre-tax cost for housing, child care, food, transportation, health care, emergency savings, and miscellaneous necessities to account for retirement savings. These family-based thresholds are then translated into living wage standards by restricting our dataset to families with one or two working parents and zero, one, or two children (for a total of six family archetypes) and taking the median value. We calculate an additional wage threshold for workers with employer-sponsored health insurance, where the family-level health care costs are set to 17% for families with no children and 29% for families with one or more children, based on average contributions identified by the Kaiser Family Foundation (KFF) in their 2023 benefits survey.

Identifying the share of struggling families

To identify the share of families in each county that struggle to make ends meet (whether or not they have a quality job, as defined later in this section), we apply our family-based cost-of-living thresholds to one-year microdata from the 2023 American Community Survey (ACS), obtained through IPUMS USA. For each family in the dataset, we append estimated annual costs-of-living data from the University of Washington based on the family's total number of infants (ages zero to two), preschool children (ages three to five), school-aged children (ages six to 12), teenagers (ages 13 to 17) and adults (ages 18-plus). For the few family-county combinations without available University of Washington data (typically those in smaller counties with an above-average number of adults or children), we linearly predict each cost value based on the county's total population, number of family members in each age group, and EPI-based cost thresholds. We then substitute the housing threshold for the actual total housing cost paid by each family (using IPUMS variables $rentgrs$ and $owncost$ for renter households and owner households, respectively), and adjust the health insurance threshold using the KFF contribution shares described above if the family indicated that they are enrolled in employer-based health insurance (via IPUMS variable $hinsemp$). Individuals are defined as being in a struggling family if the sum of these adjusted cost thresholds exceeds the family's total annual income (IPUMS variable $ftotinc$).

In supplement to this classification, we assess whether each family would struggle at a range of wage thresholds ($10 to $75 per hour). For each individual, we estimate their total number of hours worked in the reference year (the product of IPUMS variables $wkswork1$ and $uhrswork$) and total these hours by family, then determine if the family would struggle to make ends meet if they had worked that number of hours at each wage level. Importantly, while both the overall share of struggling families and the share of struggling families at each of these wage levels are useful metrics for determining who stands to benefit most from the creation of higher-quality jobs, they are not themselves indicators of job quality, since they do not account for whether each job provides benefits, nor whether workers in each family worked overtime or multiple jobs to achieve self-sufficiency.

Measuring occupational mobility

This market assessment employs a modified version of the methodology proposed in Opportunity Industries, using longitudinal survey data from the Current Population Survey (CPS) and resume records from Lightcast/Burning Glass Technologies to map the career trajectories and occupational mobility of workers in the United States. To construct our transition matrices, we filter our dataset of resumes from Lightcast to include jobs-level data for individuals who have worked in at least two jobs since January 2004, excluding records with unclassifiable occupations, inactive job records with missing end dates, and records with missing expected wage ranges, titles, or company names. Additionally, we filter out records classified within Lightcast's own proprietary specialized occupational classification taxonomy (LOT) that demonstrated highly spurious co-occurrence with occupational classifications available in the Standard Occupational Classification (SOC) system, including students, volunteers, interns, and business owners/founders.[i] These records are then sorted by individual identifier, start date, and end date, lagging occupational identifiers by row and removing records with no identifiable previous occupation, leaving a final dataset of 113.56 million recorded job transitions spanning 34.73 million workers.

While Lightcast's data is extremely comprehensive, its reliance on data from online resume profiles leads to underrepresentation of certain occupations and demographic groups relative to the nation (particularly those concentrated in lower-wage work). To counter this, we reweight our resume data based on transitions observed in the CPS's Annual Social and Economic Supplement (ASEC) data from 2004 through 2024, obtained through longitudinally linked samples from IPUMS CPS. We additionally zero-weight transitions where workers moved up or down more than a single job zone in a single step to zero, since these transitions are often a product of educational attainment gained at the time of transition and not direct career mobility.[ii]

For each occupation pair representing starting and ending occupations, we construct the transition matrix $P(i) = e^{Q \cdot i}$, representing the exponential of transition rate matrix Q at year i. For each year, the resulting probability matrix represents the probability of an individual starting in a given occupation transitioning into the ending occupation as a snapshot in time. Therefore, to measure the cumulative probability of workers making each transition between the starting year and ending year i, we calculate the cumulative probability:

$$P_{cum}(i) = \frac{1}{i}\left(e^{Q \cdot 1} + e^{Q \cdot 2} + \cdots + e^{Q \cdot i}\right),$$

Where $P_{cum}(i)$ represents the average cumulative probability of transitioning from one occupation to another by year i. Alternatively, at each year i and occupation n:

$$P(i) = e^{Q \cdot i} = \begin{bmatrix} p_{11(i)} & p_{12(i)} & \cdots & p_{1n(i)} \\ p_{21(i)} & p_{22(i)} & \cdots & p_{2n(i)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1(i)} & p_{n2(i)} & \cdots & p_{nn(i)} \end{bmatrix}, \text{ and } P_{cum}(i) = \frac{1}{i}\sum_{t=1}^{i} e^{Q \cdot t} = \begin{bmatrix} \bar{p}_{11(i)} & \bar{p}_{12(i)} & \cdots & \bar{p}_{1n(i)} \\ \bar{p}_{21(i)} & \bar{p}_{22(i)} & \cdots & \bar{p}_{2n(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{p}_{n1(i)} & \bar{p}_{n2(i)} & \cdots & \bar{p}_{nn(i)} \end{bmatrix}.$$

One limitation of these transition matrices is that they are inherently memoryless, rendering the probability of transitioning from one occupation to another at any point in time solely dependent on its proximate state, rather than other attributes that strongly influence career mobility (such as complete job history, tenure, and level of education). Despite this limitation, measurement of occupational mobility through this series of transitions remains a stronger method of mapping career paths over time than static longitudinal data.

Defining opportunity through good and promising jobs

As proposed in Opportunity Industries, this analysis uses both wage quality and occupational mobility as a basis for defining and identifying industries and occupations that provide the strongest opportunities for workers with and without four-year degrees. To identify "good jobs," we take the insurance-based living wage standard identified earlier in this appendix and, using percentile earnings data from Lightcast, use linear interpolation to estimate the total share of jobs in each occupation that pay an hourly wage greater than or equal to that threshold. For each percentile:

$$\hat{p}_w = 1 - \left(pctile_{lb} + \frac{lwage - pwage_{lb}}{pwage_{ub} - pwage_{lb}} \cdot (pctile_{ub} - pctile_{lb})\right),$$

Where $pctile$ represents the percentile value, $pwage$ represents the wage value at each percentile, $lwage$ represents the living wage threshold (assuming employer-sponsored health insurance is provided), and ub and lb represent the use of upper-bound and lower-bound values for each percentile range, respectively. For this calculation, we use 10th, 25th, 50th, 75th, and 90th percentile wages provided by Lightcast, and estimate first and

99th percentile values based on the ratio distance between 10th/25th and 75th/90th percentile values for each occupation. While these additional percentiles may not perfectly align with each occupation's observed wage distribution, they are helpful in accounting for edge cases in the data, and provide a conservative approach for limiting the share of occupations predicted to have exactly zero "living-wage jobs," and vice versa.

Beyond meeting this wage threshold, this analysis requires jobs to provide employer-sponsored health insurance to qualify as a "good job," as employer health insurance is a helpful proxy for other elements of job quality (such as the availability of other benefits and employment stability). To calculate the expected share of jobs that provide employer health insurance, we use national data from the CPS ASEC's health insurance supplement to calculate the share of workers in each occupation from 2004 to 2024 with an employer that offered group health insurance and were eligible for enrollment (whether or not they ultimately chose to do so). We then weighted these national estimates using county-level data from the 2023 ACS one-year sample based on the share of workers in each occupation that were covered by employer health insurance at the time they were surveyed. These county-level insurance shares were multiplied by the share of living-wage jobs in each occupation, providing a final estimated share of total jobs that both paid a living wage and provided employer health insurance in 2023.

From this total share of "good jobs" in each occupation and county, we further predicted the share of "promising jobs" that provide strong mobility pathways to good jobs within 10 years, even if they do not qualify as good jobs now. To calculate the share of promising jobs, we multiplied the national cumulative transition probability for each occupation by each percentile wage threshold and health insurance share, and summed these predicted values to provide a total expected wage rate and health insurance accessibility rate for that occupation over 10 years. Promising jobs were then estimated based on linear interpolation of wage and health insurance accessibility at 10 years for each occupation as a share of jobs in the starting year and occupations *not* classified as "good."

Qualitative assessment

Brookings undertook a series of interviews, roundtables, and other stakeholder interactions from fall 2024 to spring 2025 to interpret data analysis and findings with qualitative insights. These activities sought to contextualize the analysis and inform sector selection with market intelligence around industry challenges and opportunities, business strategy, civic dynamics, policy issues, and other trends unavailable from quantitative data. These activities also served the broader stakeholder engagement objective to engage diverse input to build further regional support and buy-in for the effort. In total, these contacts touched approximately 90 stakeholders through interviews and roundtables, and dozens more through substantive engagement in presentations, events, and conferences.

Targets for engagement included:

- Individual firms, sector groups, and investors in major prioritized (or likely to be prioritized) traded industries, who addressed factors impacting industry performance (both unique to the region and reflecting broader macroeconomic, industry, and policy dynamics); talent needs and challenges; human resources and job quality practices; and the effectiveness of and outstanding needs from regional, local, and state economic and workforce development service delivery.

- Participants in the region's innovation, commercialization, and high-growth entrepreneurship ecosystem, who addressed dynamics influencing the strength and density of the region's innovation and high-growth entrepreneurship ecosystem (e.g., capital access, tech transfer/commercialization capabilities); views of market potential of sectors; the effectiveness of and gaps within existing interventions; and distribution and reach across the region.

- Workforce development actors (e.g., workforce development boards, community colleges), who addressed talent systems' approach to industry engagement and prioritization; definitions of job quality and approach for promotion; and alignment with economic development systems and institutions.

Brookings partnered with representatives of the Southeastern Pennsylvania Economic Collaborative (which includes leaders from Bucks, Chester, Delaware, Montgomery, and Philadelphia counties and other regional organizations, including the Chamber of Commerce for Greater Philadelphia, Visit Philadelphia, and the Delaware Valley Regional Planning Commission) to identify and engage stakeholders in these categories across their jurisdictions, recognizing local relationships and market perspectives. Interviews were typically conducted via video conference, largely over the first quarter of 2025. A standard set of questions was adapted for each session to reflect individuals' backgrounds. Brookings also partnered with the Chester County Economic Development Council to convene a series of four industry roundtables in Exton, Penn., in February 2025.

In tandem with interviews and roundtables, Brookings used numerous presentations to stakeholder groups to test and refine findings and messages. These groups included the Southeastern Pennsylvania Economic Collaborative, the Pew Charitable Trusts' Roadmap for Quality Jobs Steering Committee, the Chamber of Commerce for Greater Philadelphia's CEO Council for Growth and Select Greater Philadelphia Advisory Board, the regional Workforce Development Board directors, the Philadelphia Works board of directors, and the Federal Reserve Bank of Philadelphia's Economic Mobility Summit. Brookings also attended events such as the PACT Phorum Technology Conference to gather market insights and engage with relevant stakeholders.

Industry clustering and prioritization

The strategic industries described in this market assessment were identified through a cluster analysis of Southeastern Pennsylvania's economy, using firm-level descriptive data from Crunchbase, occupation-to-industry staffing matrices from Lightcast, and regional input-output tables to identify supply chain linkages and firm interrelatedness between industries across Southeastern Pennsylvania. Inputs from these data sources were combined into a distance matrix and analyzed through k-means clustering, resulting in more than 250 clusters linked together through common supply chains, firm activities, and talent pools.

The industry clusters identified through this analysis were then analyzed through the labor market and job quality analyses detailed above to gauge the relative economic strength, competitiveness, and opportunity profile provided by each cluster. Industries were prioritized for intervention based on their relative scale, competitiveness, and concentration/intensity (measured by location quotient, with adjustments for regional effects). To address the central concern of economic opportunity and mobility in Southeastern Pennsylvania, clusters were only selected for intervention if at least half of all jobs within them qualified as good or promising, as well as having above-average mobility for workers without a four-year degree within those opportunity jobs. Priority clusters were further vetted against market intelligence developed via qualitative research and stakeholder engagement to determine areas with the strongest opportunities for impact.

---

[i] Business owners and founders are often categorized as chief executive officers across Lightcast datasets, which include labor market information for workers not traditionally counted in the OES and QCEW. These workers demonstrate highly different career pathways and wage trajectories than traditional firm-based chief executive officers. Because these distinguishing characteristics cannot be captured by the SOC-5 taxonomy employed in this analysis, job transitions into this category appear to offer much stronger wage and mobility pathways for workers than they realistically represent. The removal of these records limits the appearance of strong economic mobility for workers who have transitioned out of the traditional workforce for entrepreneurial reasons, rather than achieving mobility through career growth.

[ii] Job zones are based on educational attainment data from Lightcast, categorizing the level of educational attainment typically required for entry into each occupation. In this assessment, all jobs that do not require a bachelor's degree for entry are considered to be part of the same job zone, while occupations requiring a bachelor's degree, master's degree, or doctoral degree are categorized into their own distinct respective job zones. Therefore, this analysis zero-weights transitions observed between workers from occupations with no degree requirements into occupations that typically require a postgraduate degree, as well as transitions between workers from occupations that typically require a bachelor's degree into occupations that require a doctoral degree (and vice versa).