Understanding Self-Employment Trajectories and their Effects on Wellbeing in the United States

Joelle Abramowitz

Andrew Joung

Extended Abstract

Approximately 20 percent of workers report engaging in some self-employment activity. However, as discussed in the National Academies of Sciences, Engineering, and Medicine (2020) report on measuring alternative work arrangements, existing data sources on these kinds of work arrangements have limitations in their ability to be used to understand the changing nature of work and to develop appropriate policy. These limitations include a dearth of data identifying heterogeneity in the nature of these work arrangements and how workers move across them over their lives. For example, an individual pursuing self-employment in the transportation sector could choose to innovate a new platform or technology, drive for an app-based ride-sharing service, advertise their own chauffeur services, drive on a contract basis for an established business, or manage their own or someone else's established business. The characteristics of these jobs - the barriers to entry, risks, work stresses, and compensation - are likely to vary considerably. However, in most existing data sources, we would be unable to meaningfully differentiate these jobs. Understanding heterogeneity in self-employment trajectories and their effects on wellbeing has become all the more important as new technologies such as electronic platforms have introduced new means of engaging in self-employment with potentially more far-reaching effects on the economy.

It is important to identify these distinctions in roles to understand patterns of entry and exit from self-employment and understand the relationship between these different kinds of work and wellbeing over the life course. For example, what types of arrangements would workers engage in if they were not doing electronic platform work? Does electronic platform work induce workers to engage in other self-employment and entrepreneurial activities? Or does it serve as a less stressful option for older workers transitioning to retirement?

This paper fills existing gaps by using novel data from two large-scale and long-running surveys, the Panel Study of Income Dynamics (PSID) and the Health and Retirement Study (HRS), to examine trajectories in self-employment work arrangements, identifying separately business ownership, platform gig work, informal self-employment, and formal self-employment. The approach uses novel data constructed using machine learning and internal respondent narratives on industry and type of work and employer names collected in the 2003-2019 PSID and 1994-2020 HRS in conjunction with public PSID and HRS data. The paper uses these data to identify: (1) how workers transition across different work arrangements across survey waves and at retirement, (2) patterns in cumulative exposure to different types of work arrangements, and (3) patterns in sequences of exposure to different types of work arrangements. The paper then examines how each of these trajectories affects emotional, physical, and financial wellbeing.

1 Measuring Self-Employment

Self-employment is notoriously hard to measure. While both administrative data and surveys are able to provide some insight about self-employment, they each face challenges. As a result of these challenges, discrepancies appear across administrative and survey data sources in identifying trends in self-employment broadly and in specific arrangements such as contingent work and gig employment (Abraham et al., 2018, 2021; Allard and Polivka, 2018; Jackson et al., 2017; Katz and Krueger, 2019).

Administrative data are derived from tax reporting and can be used to identify wage and salaried employment separately from self-employment based on the types of income reported in tax filings. However, third-party reports by firms suffer from incomplete coverage (Garin et al., 2022). In the absence of third-party reporting, taxpayers report self-employment income through their tax filings on Schedules C and SE. However, taxpayers face incentives to strategically report taxable income, leading to concerns regarding the accuracy of administrative data (Mortenson and Whitten, 2020; Saez, 2002) with the IRS estimating only a 44% tax reporting compliance rate among self-employed workers (Slemrod, 2016).

Since individuals do not face the same economic incentives to misreport on surveys as on their tax filings, surveys provide a valuable complement to administrative data, especially in the context of measuring self-employment. However, surveys were designed to capture traditional wage and salaried employment, and accordingly, an ample literature documents myriad issues as survey questions may not capture the way individuals perceive their work, especially in less traditional work arrangements (Allard and Polivka, 2018; Bureau of Labor Statistics, 2018; Abraham and Amaya, 2019; Bracha and Burke, 2021; Abraham et al., 2023), and can produce divergent estimates based on sampling modes, methods, and timing (Katz and Krueger, 2019).

The present study adds to the literature by exploring heterogeneity in self-employment work arrangements based on respondents' descriptions of their work responsibilities and employer names. By using PSID and HRS data, this work benefits from a high survey response rate as well as the breadth of information collected in the PSID and HRS both contemporaneously and longitudinally.

2 Data and Methods

This analysis uses the 2003-2019 PSID and the 1994-2020 HRS. The PSID is used to consider self-employment for workers of all ages while the HRS is used to focus on outcomes for older workers during and after their transition to retirement. Both surveys are valuable for this analysis as they are nationally-representative and longitudinal, fielded every two years, and include questions on a breadth of topics including employment, income, and physical and mental health.

The PSID is a longitudinal dataset that began in 1968 with a sample of approximately 5,000 U.S. households; it was updated annually through 1997 and bi-annually thereafter. As of 2017, it had grown to include over 10,000 families and 24,000 individuals. In addition to publicly-available PSID data, our analysis leverages internal data collected on narrative descriptions of industry and occupation and employer names to classify work arrangements

into a useful framework. The PSID collects this information for all of the respondent's and the spouse's jobs held since January 1 of the prior survey year through the time of the survey.

The HRS is a longitudinal survey of a representative sample of approximately 20,000 Americans over age 50 and their spouses, updated every two years with new cohorts added every six. Similar to the PSID, in addition to publicly-available HRS data, our analysis leverages internal data collected on narrative descriptions of industry and occupation to classify work arrangements into a useful framework. The HRS collects this information for the respondent's and the spouse's current main jobs held at the time of the survey.

The project makes use of Abramowitz et al.'s (2023) classification of work arrangements for the 2003-2019 waves of the PSID and Abramowitz and Kim's (2021) classification of work arrangements for the 1994-2020 waves of the HRS, whereby employer names (PSID only) and narrative responses (PSID and HRS) to the open-ended industry and type of work questions were each coded as one of five work arrangements (platform-mediated gig work, informal self-employment, formal self-employment, business owners, wage and salaried employees) and a small number assigned no label due to insufficient information. We used machine learning to automate the classification. Two reviewers classified the same subset of 30% the data according to the described schema, with disagreements adjudicated by a third reviewer, to be used to train a machine learning model to classify of the remainder of the data. Reviewers also classified records for which the model did not confidently predict a classification, following the same procedure as for producing the training data.

3 Preliminary Findings

3.1 Wave-to Wave Transitions in Work Arrangements in the PSID

In preliminary work, we first document how workers transition across work arrangements from one survey wave to the next in the PSID. Table 1 presents transition matrices across four work arrangements (wage and salaried employment, informal self-employment, formal self-employment, and business ownership) and non-employment. To observe changes in transitions over time, we report separate wave-to-wave transitions for 2003-2009 and 2011-2019. The work arrangement in the prior survey wave is represented vertically while the survey wave work arrangement is represented horizontally. Each cell shows the percentage of respondents having that combination of prior and current work arrangements.

Table 1 shows that across all work arrangements, respondents are most likely to persist in the work status that they had in the prior survey wave. However, relative to non-employment and wage and salaried employment, self-employment is associated with greater diversity in transitions. Focusing first on 2003-2009 in Panel A, we see that 74.0% of the non-employed and 85.3% of wage and salaried workers in the previous wave stay in their respective roles. On the other hand, we see that self-employed workers are not nearly as likely to remain in their roles across survey waves with only 36.9% of informally self-employed workers and 37.4% of formally self-employed workers staying in their roles

across waves. Business owners are also less likely to persist in their roles than wage and salaried employees, but more likely to remain in their roles than any other type of self-employment, with 49.8% of business owners remaining in their roles across survey waves. In addition, we see that relative to all other work arrangements, the informally self-employed are more likely to enter non-employment: 20.2% of those informally self-employed in the previous wave became non-employed in the current survey wave, nearly double the probability of any other work arrangement.

Examining transitions from 2011-2019 in Panel B, we see largely similar patterns to those that we document from 2003-2009. One notable change is a rise in persistence within work arrangements. For informally self-employed workers, we see that the probability of staying informally self-employed across waves increases from 36.9% in 2003-2009 to 47.1% in 2011-2019. For business owners, we see a larger rise with the probability of remaining a business owner increasing from 49.8% to 70.2%.

3.2 Transitions to Retirement in the HRS

To understand how the transition to retirement is driven by changes in self-employment work arrangements, Table 2 presents a transition matrix across wage and salaried employment, three self-employment arrangements (informal and formal self-employment, business management, and business ownership), and non-employment for respondents who appear in the HRS both pre- and post-retirement. Work arrangements are identified as the modal arrangement in each of the pre- and post-retirement periods. The pre-retirement arrangements are represented vertically while the post-retirement arrangements are represented horizontally. Panel A presents the percentage of each pre-retirement arrangement for each transition, and Panel B presents the number of respondents for each transition.

Panel A shows that the majority of employees and those not working pre-retirement transition to not working post-retirement. In contrast, most engaged in informal or formal self-employment pre-retirement continue to do so post-retirement. The majority of managers transition to informal or formal self-employment or become employees while owners tend to stay in their ownership roles or transition to informal or formal self-employment.

Observing the magnitudes of these transitions in Panel B sheds light on the finding that the share of workers engaged in self-employment increases with age. While only 7.9 percent of employees transition to informal or formal self-employment, they represent many more individuals than the 13.1 percent of the informal or formal self-employed (and 27.2 percent of managers and 9.5 percent of owners) who transition to wage and salaried employment.

4 Discussion

For researchers and policymakers, our preliminary results emphasize the importance of capturing the heterogeneity within self-employment to understand its changing nature and future trajectory. Prior work has shown that administrative data miss substantial amounts of self-employment activity at both the intensive and extensive margins and surveys generally do not probe to the necessary extent to identify work arrangements of interest. Using novel

data, we are able to identify these self-employment work arrangements and distinguish how workers more across different types of arrangements in distinctive ways over the life course and at retirement. The paper will build on these preliminary findings to examine patterns in cumulative exposure and sequences of exposure to different types of work arrangements and examine how they affect emotional, physical, and financial wellbeing.

References

Abraham, Katharine G., and Ashley Amaya. 2019. 'Probing for Informal Work Activity.' Journal of Official Statistics 35(3):487–508.

Abraham, Katharine G., John Haltiwanger, Claire Hou, Kristin Sandusky, and James R. Spletzer. 2021. 'Reconciling Survey and Administrative Measures of Self-Employment.' Journal of Labor Economics 39(4):825-860.

Abraham, Katharine, John Haltiwanger, Kristin Sandusky, and James Spletzer. 2018. 'Measuring the Gig Economy: Current Knowledge and Open Issues.' Working Paper No. 24950. Cambridge, MA: National Bureau of Economic Research.

Abraham, Katharine G., Brad Hershbein, Susan N. Houseman, and Beth Truesdale. 2023. 'The Independent Contractor Workforce: New Evidence on its Size and Composition and Ways to Improve its Measurement in Household Surveys.' Working Paper No. 30997. Cambridge, MA: National Bureau of Economic Research.

Abramowitz, Joelle, Jenna Kim, and Jinseok Kim. 2023. "C-NARS/P: An Open-Source Tool for Classification of Narratives in Survey Data Using Pre-Trained Language Models." Technical Report.

Abramowitz, Joelle, and Jinseok Kim. 2021. "C-NARS: An Open-Source Tool for Classification of Narratives in Survey Data." University of Michigan Technical Report.

Allard, Mary D., and Anne E. Polivka. 2018. 'Measuring Labor Market Activity Today: Are the Words Work and Job Too Limiting for Surveys?' Monthly Labor Review. U.S. Bureau of Labor Statistics.

Garin, Andrew, Emilie Jackson, and Dmitri Koustas. 2022. 'Is Gig Work Changing the Labor Market? Key Lessons from Tax Data.' National Tax Journal 75(4):791-816.

Jackson, Emilie, Adam Looney, and Shanthi P. Ramnath. 2017. 'The Rise of Alternative Work Arrangements: Evidence and Implications for Tax Filing and Benefit Coverage.' Working Paper No. 114. Internal Revenue Service Office of Tax Analysis.

Katz, Lawrence F., and Alan B. Krueger. 2019. 'Understanding Trends in Alternative Work Arrangements in the United States.' Working Paper No. 25425. Cambridge, MA: National Bureau of Economic Research.1

Mortenson, Jacob A., and Andrew Whitten. 2020. 'Bunching to Maximize Tax Credits: Evidence from Kinks in the US Tax Schedule.' American Economic Journal: Economic Policy 12(3):402-32.

National Academies of Sciences, Engineering, and Medicine. 2020. 'Measuring Alternative Work Arrangements for Research and Policy.' Washington, DC: The National Academies Press.

Saez, Emmanuel. 2002. Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses. The Quarterly Journal of Economics 117(3):1039-1073.

Slemrod, Joel. 2016. 'Caveats to the Research Use of Tax-Return Administrative Data.' National Tax Journal 69(4):1003-1020.

Table 1: Wave-to-Wave Employment Transitions in the PSID

Panel A: Wave-to-Wave Transitions in Employment Status (2003-2009)						
		Current Wave				
		Not Working	W&S Employment	Informal SE	Formal SE	Business Ownership
Prior Wave	Not Working	74.0%	21.2%	2.9%	1.1%	0.8%
	W&S Employment	10.5%	85.3%	1.6%	1.6%	1.0%
	Informal SE	20.2%	22.6%	36.9%	11.6%	8.7%
	Formal SE	7.9%	26.2%	11.9%	37.4%	16.6%
	Business Ownership	7.1%	16.8%	9.4%	17.0%	49.8%

Panel B: Wave-to-Wave Transitions in Employment Status (2011-2019)						
		Current Wave				
		Not Working	W&S Employment	Informal SE	Formal SE	Business Ownership
Prior Wave	Not Working	82.2%	13.3%	2.7%	1.0%	0.8%
	W&S Employment	10.5%	86.5%	1.0%	1.0%	0.9%
	Informal SE	22.3%	17.2%	47.1%	8.1%	5.3%
	Formal SE	10.7%	24.3%	9.3%	43.4%	12.3%
	Business Ownership	8.5%	11.4%	4.0%	5.8%	70.2%

Source: Internal PSID narrative data on industry and occupation and employer names (2003-2019) classified into work arrangement types. Additional information on work status comes from the public PSID (2003-2019) merged to the narrative data classified into work arrangement types.
We report the probability of a respondent transitioning to a given current main job type conditional on their main job type in the prior survey wave. Estimates use cross-sectional PSID weights.
Abbreviations: W&S, wage and salaried; SE, self-employment.

Table 2: Retirement Transitions by Pre-Retirement Work Arrangement in the HRS

Panel A: % of Pre-Retirement Work Arrangement

	Post-Retirement				
Pre-Retirement	Not Working	W&S Employment	Informal or Formal SE	Business Mgmt.	Business Ownership
Not Working	87.8%	8.3%	3.8%	0.0%	0.0%
W&S Employment	53.5%	38.2%	7.9%	0.2%	0.2%
Informal or Formal SE	31.1%	13.1%	51.6%	1.2%	3.1%
Business Mgmt.	31.4%	27.2%	33.6%	6.9%	0.9%
Business Ownership	31.9%	9.5%	35.4%	2.0%	21.2%

Panel B: Number of Respondents

	Post-Retirement				
Pre-Retirement	Not Working	W&S Employment	Informal or Formal SE	Business Mgmt.	Business Ownership
Not Working	12,609	1,196	543	0	6
W&S Employment	46,690	33,340	6,887	162	214
Informal or Formal SE	4,566	1,929	7,578	172	450
Business Mgmt.	177	153	189	39	5
Business Ownership	319	95	354	20	212

Source: 2018 RAND HRS Longitudinal File and 1994-2018 HRS self-employment industry and occupation narratives.

Abbreviations: W&S, wage and salaried; SE, self-employment.