The Causes and Consequences of Self-Employment over the Life Cycle

John Eric Humphries*
Yale University

February 12, 2024

Abstract

This paper uses population panel data from Sweden to investigate the causes and consequences of self-employment over the life cycle, and to evaluate how self-employment decisions can be influenced by policy. In the first part of the paper, I use machine learning methods to summarize the patterns of self-employment behavior observed in the data. I find that careers involving self-employment fit into a small number of economically distinct groups. Some self-employment spells are short, with minimal capital investment and rapid return to paid employment, while others persist and have substantial capital devoted to the business from the outset. Guided by these descriptive results, I develop and estimate a dynamic Roy model in which self-employment decisions depend on factors such as cognitive and non-cognitive skills, prior work experience, the cost of capital, and other labor market opportunities. The model integrates traditional models of dynamic career choice, which feature human capital investment, with models of business start-up, which feature physical capital investment. I estimate the model and use it to evaluate policies designed to promote self-employment. Cognitive and non-cognitive skills, education, and prior work experience are important determinants of the types of businesses individuals start, how much capital they employ, and how long they remain in self-employment. Subsidies that incentivize self-employment are generally ineffective in promoting long-lasting firms and in improving the welfare and earnings of those induced to enter self-employment.

*I gratefully acknowledge financial support from the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond) grant P12-0968 that enabled the data collection for this project. I would like to thank Stephane Bonhomme, James J. Heckman, Erik Hurst, Juanna Joensen, Scott Kominers, Jim Marrone, Costas Meghir, Magne Mogstad, Derek Neal, Karl Schurter, Winnie van Dijk, and Gregory Veramendi for helpful feedback and support.

1 Introduction

More than 10% of working men in the United States were self-employed in 2015. The new businesses started by the self-employed are often seen as an important source of growth and innovation for the economy.[1] As a result, governments seek to promote small business creation through tax breaks, grants, and training programs. Yet, the self-employed are a heterogeneous group, and policies may be increasing the number of self-employed without promoting growth and innovation. Indeed, many self-employment spells are short, most do not involve substantial investments, and among the businesses that last, many never grow.[2]

This paper uses panel data on workers and firms from Sweden to investigate the causes and consequences of self-employment over the life cycle, and evaluate how self-employment decisions can be influenced by policy.[3] To do this, the paper proceeds in two parts. In the first part, I apply machine learning methods to isolate distinct patterns of self-employment behavior over the life cycle and present an intuitive two-period model to interpret these results. Guided and motivated by this descriptive evidence, the second part of the paper develops a dynamic model in which life cycle factors such as cognitive and non-cognitive skills, prior work experience, the cost of capital, and labor market opportunities explain these patterns of self-employment. I use the estimated model to quantify the importance of skills, career histories, and the cost of capital on self-employment decisions, and to evaluate the consequences of policies designed to promote self-employment.

The first part of the paper presents three sets of results. Using machine learning methods for clustering discrete time-series, I document the variety of patterns of self-employment (SE) behavior over the life cycle and develop a data-driven taxonomy of these distinct profiles. I find that careers involving self-employment fit into one of seven distinct groups: (1) late-onset unincorporated SE, (2) full-career unincorporated SE, (3) late-onset incorporated SE, (4) full-career incorporated SE, (5) low labor force participation with occasional spells of SE, (6)

[1]For example, Haltiwanger et al. (2013) documents that young firms contribute substantially to gross and net job creation. In particular, new start-ups account for 3% of employment but 20% of gross job creation. Haltiwanger et al. (2016) further shows that young firms disproportionately contribute to growth in output. See Decker et al. (2014) for a review of this literature.

[2] Schoar (2010), Hurst and Pugsley (2011), Haltiwanger et al. (2013) and Levine and Rubinstein (2013) all also emphasize the heterogeneous nature of self-employment.

[3]The data contains detailed longitudinal information on the individuals' background and career choices, industry characteristics, and firm characteristics for the full population. In particular, I link self-employed individuals to detailed information on the businesses they create, including the amount of capital devoted to the business.

paid employment with a brief spell of self-employment, and (7) prolonged schooling followed by late-onset SE.[4] I show that there are systematic differences in who selects into each self-employment group, and what they do while self-employed. Some groups have above average cognitive and non-cognitive skills and earn more than those in paid employment, while other groups have below average skills and earnings. Moreover, the various groups use substantially different amount of capital and labor upon first entering self-employment.

I develop an intuitive two-period model to rationalize and interpret the patterns of self-employment decisions observed in the data. In the two-period model, agents choose between paid employment and self-employment to maximize their expected discounted utility. Agents enter self-employment based on comparative advantage that depends on ability and career-specific experience. If agents choose to be self-employed, they also choose how much capital they devote to their business. I assume that agents can rent capital at a fixed market rate, but that there are spell-specific capital adjustment costs associated with capital market frictions and the costs of doing business.[5] In the model, agents may choose to self-employ to gain additional human capital. When self-employed, agents who do not expect to remain self-employed have incentives to start smaller, less capital-intensive businesses. I use the model to interpret new empirical evidence on how career dynamics affect agents' entry, exit, and investment decisions and show that the empirical evidence is consistent with the implications of the simple two-period model.

Building on the two-period model, the second part of the paper develops and estimates a structural model in order to both quantify the importance of abilities and career histories on self-employment decisions and to evaluate the impacts of counterfactual policies. The structural model acknowledges that self-employment decisions are made in a larger labor market context where heterogeneous individuals aim to maximize their well-being over the life cycle, invest in human capital, and plan for the future. The model integrates traditional models of dynamic career choice that feature human capital investment with models of business start-up that

[4] Separating incorporated and unincorporated self-employment is consistent with Levine and Rubinstein (2013) who document key differences in who start incorporated and unincorporated businesses and what they do while self-employed.

[5] This builds on the work of Evans and Jovanovic (1989) and Buera (2009) by considering a dynamic self-employment decision with human capital. Unlike many classic models of entrepreneurship, such as Lucas (1978), Evans and Jovanovic (1989), or Murphy et al. (1991), individuals from across the distribution of ability may select into self-employment.

feature costly capital investment. In the model, agents choose between self-employment and paid employment in either the white-collar or blue-collar sector, where "career" is the choice of sector and organizational form. Agents acquire career-specific human capital that may also be valued in other careers. When self-employed, agents choose the amount of capital to employ in their business.

Using the structural model, I demonstrate that, for the most part, policies designed to promote self-employment tend to induce small and short-lived businesses. On average, those induced into self-employment by the subsidies earn substantially less in self-employment than in their previous careers. Some high-ability individuals start large, lasting businesses and have higher lifetime earnings but only small welfare gains. Those induced out of non-employment also have higher lifetime earnings through increased labor force attachment. Finally, I show that policies boosting cognitive or non-cognitive skills increase the overall amount of self-employment, but across different margins.[6] Boosting cognitive skills increases participation in all forms of white-collar labor, while boosting non-cognitive skills raises both blue-collar and white-collar self-employment, primarily by reducing non-employment.

I also use my structural model to evaluate how the impacts of policies differ based on the age of those targeted. I find that policies targeting younger individuals are preferable to policies targeting older individuals for two reasons. First, policies targeting those early in their careers tend to induce more people into self-employment and cause a smaller initial drop in earnings upon entering self-employment because individuals have developed less human capital in paid employment. Second, the gains to self-employment take several years to be fully realized, giving younger individuals more time to reap potential benefits.

This paper contributes to our understanding of policies designed to promote self-employment. Many governments promote policies that encourage self-employment and small business ownership. In the United States, the tax code has traditionally favored small businesses and, through the Small Business Administration, the government fosters small businesses through grants, subsidized loans, and the provision of information. Similar policies exist in many European countries. For example, Sweden spent approximately 1% of its GDP promoting small businesses and entrepreneurship through a combination of tax breaks, grants, and educational programs in

[6] See Heckman and Kautz (2014) for a review of interventions shown to boost cognitive and non-cognitive skills.

2009 (Vikstrom, 2011).[7] To properly evaluate policies designed to promote self-employment, we must understand how different policies may induce people with different backgrounds and career profiles to enter self-employment.[8] Using the model developed below, I evaluate a number of possible counterfactual policies, and show that they impact a heterogenous group of individuals. To my knowledge, this is the first paper to estimate the long-run welfare and wage impacts of self-employment policies.

Contribution to the literature

This paper contributes to the literature on self-employment in several ways. First, little consensus has been reached on what causes people to enter self-employment. In a review of the literature, Parker (2009) finds that age, experience, wealth, and having entrepreneurial parents are associated with being self-employed but finds no consensus on many other factors.[9] Where associations exist, there is little evidence on the underlying mechanisms. Many papers explicitly model the choice to start a business, but these are primarily static models in which agents choose between entrepreneurship and paid employment in order to maximize profits.[10] By contrast, this paper provides the first unifying framework to compare the relative importance of the various factors believed to affect self-employment decisions and to evaluate how these decisions might vary across individuals and over the life cycle.

Second, there is little research examining how self-employment affects current and future earnings. Early work by Evans and Leighton (1989) documents that self-employment has lower labor market returns than paid employment. This is supported by Bruce and Schuetze (2004a), who find that brief self-employment spells appear to reduce wages when agents return to paid employment. Hamilton (2000) similarly documents that the self-employed earn less than their

[7] A number of papers use reduced-form methods to evaluate policies designed to promote self-employment. See, for example, Elert et al. (2015), Goetz et al. (2010), Bennmarker et al. (2009), Bruce and Schuetze (2004b), Benus et al. (2008).

[8] For evaluation, we may wish to study how many large or long-lasting businesses are created, but we would also like to understand the opportunity cost of inducing individuals into self-employment (i.e. what they left behind in order to self-employ).

[9] For example, Parker (2009) documents 69 studies that show education is a positive determinant of entrepreneurship, 21 studies that find education is a negative determinant, and 27 that find no correlations to education. See Simoes et al. (2015) for a more recent review which reaches similar conclusions. Also see Blanchflower and Oswald (1998), Hamilton (2000), and Hamilton et al. (2015).

[10] See, for example, Lucas (1978), Evans and Jovanovic (1989), and Murphy et al. (1991).

employed counterparts despite the risky nature of self-employment.[11] Hamilton argues that this can be explained by large non-pecuniary benefits to self-employment. A small number of papers consider dynamic models of entrepreneurial decisions. Starting with Jovanovic (1982), most of these papers involve agents learning about their ability as entrepreneurs through experimentation in entrepreneurship.[12] Overall, there is little research on how self-employment choices fit into the broader career decisions of individuals. This paper addresses this question by building a framework that nests self-employment decisions within a broader model of sequential career choice. The model allows me to simulate detailed counterfactual earnings while taking into account selection into self-employment. Thus, I am able to study the returns to self-employment over the life cycle and across individuals.

Third, the literature on productivity differences among firms and firm dynamics does not typically consider the interplay between the life cycle dynamics of the individual and earnings dynamics and entry and exit decisions of the businesses they create. Empirical work such as Haltiwanger et al. (2013) and Decker et al. (2014) typically have little information about the founders of new businesses and why they entered self-employment.[13] In this paper, I document that, for most new businesses, the most important assets are the founder's skills and experience. As a result, a business's performance depends on its founder. Moreover, entry and exit decisions are closely tied to labor market dynamics and opportunity costs of individuals. For example, some people enter self-employment to start lasting businesses, while others may self-employ after a labor market shock or may self-employ as a form of marginal labor force participation. I show that individual characteristics and career dynamics of an individual are highly predictive of how their firm performs and if it lasts. This is important for understanding how often firms "fail" versus how often an individual's other labor market opportunities move them away from self-employment.

Fourth, a growing literature argues that many small businesses do not fit classic models of

[11] See Hurst and Pugsley (2016) and Hamilton et al. (2015) for two other static models of self-employment decisions that include non-pecuniary benefits.

[12] This work has been extended by Manso (2016), who demonstrates that in such learning models cross-sectional analysis of the returns to entrepreneurship will be biased and by Dillon and Stanton (2016), who explicitly model how the option value to learning raises the overall return to self-employment and how briefly trying out self-employment affects wages upon returning to the employed sector. Hincapi (2020) additionally builds a dynamic model of entrepreneurship with correlated learning and human capital accumulation.

[13] Some studies specifically model individuals' or firms' decisions to enter and exit new businesses, but these models typically assume that all new self-employed individuals aim to innovate or grow. See, for example, Jovanovic (1982), Ericson and Pakes (1995), and Dillon and Stanton (2016).

business start-up. For example, Levine and Rubinstein (2013) argue that there are fundamental differences between those who choose to start unincorporated businesses and those who choose to start incorporated businesses. Hurst and Pugsley (2011) document that many new small businesses have no intention of growing, yet persist. Block and Sandner (2009) argue that the self-employed can be split into "necessity" entrepreneurs who self-employ due to lack of other options and "opportunity" entrepreneurs who seek to bring new ideas to the market or take advantage of other market opportunities. This paper provides a framework in which individuals can start new growth-oriented businesses but which also permits for various other forms of self-employment described above. Moreover, I can document what background characteristics, skills, and career histories lead to the various self-employment behaviors.

This paper proceeds as follows. Section 2 describes the data. Section 3 documents new empirical evidence on self-employment behavior over the life-cycle. Section 4 develops the structural model. Section 5 discusses estimation. Section 6 presents the results and considers policy counterfactuals. Section 7 concludes.

2 Data

This paper combines several sources of Swedish administrative data. The paper focuses on all Swedish men born between 1965 and 1977 and studies labor market outcomes between 1990 and 2013.[14] This analysis uses annual-level data from the longitudinal integration database for health insurance and labor market studies (LISA), a linked register database maintained by Statistics Sweden which extends previous versions of the registry data used for labor market research (LINDA and LOUISE).[15] The data combines and links a number of Swedish registers. Specifically, the registers collect data on "labor market, educational and social sectors." The data focus on individuals but provide links with family members and employers. This paper links and uses two different portions of the LISA data: (1) Individual records which provide detailed accounts of income, employment, and education and (2) company-level data on the number of employees, current assets, fixed assets, and industry. The LISA data is then linked to military enlistment records.

[14]This paper focuses on native Swedish men as measures of skills come from the military enlistment exams are not available for the majority of women and immigrants.

[15]See http://www.scb.se/lisa-en for additional details.

Data on career choices and self-employment: Self-employment classifications are based on the individual's primary source of labor market income during the year. The data comes from tax records submitted by employers and tax records on individual declared business income. An alternative classification for self-employment based on primary work activity in November was also constructed and aligned closely with my primary classification. Two-digit 2002 SNI industry codes were used to construct industry. Crosswalks between the 2002 and 2007 codes were constructed using businesses that were present in both 2006 and 2007. A similar approach was used to link 1992 and 1969 industry codes in order to construct consistent industry categories over time.

Background characteristics: By linking the sample to their parents, I construct measures of parental education, average parental wealth, average parental income, and whether either parent reported being self-employed between 1990 and 2000. Using the LISA data, I additionally construct indicators for region of residence, educational attainment, and area of study.

Measures of cognitive and non-cognitive skill: Every male Swedish citizen in the birth cohorts studied had to complete conscripted military service. To aid in assigning individuals to tasks, all males are required to attend a day of military testing. For each male in the sample, I have (1) measures of cognitive ability and (2) evaluation by military psychologists designed to assess the ability of the individual to handle the rigors of active military duty.[16] The military psychologists gives each person an overall score which, based on verbal accounts, aims to measure a combination of stress tolerance, conscientiousness, and propensity for leadership.[17] These scores provide baseline measures of cognitive ability and non-cognitive ability for the sample, allowing me to control for detailed baseline heterogeneity that is usually not possible using register data.

In summary, the data provides a wealth of novel information on individuals, their career choices, and the businesses they start while self-employed.[18] To my knowledge, this is the first paper to link register data between firms and individuals to study self-employment or

[16]For the birth cohorts studied in this paper, all men had mandatory military service with few exceptions. As a result, individuals had no incentives to distort their evaluations.

[17]See Lindqvist and Vestman (2011) for a detailed overview of the measures collected during the military exams.

[18]As laid out in Section C of the appendix, the business environments in the United States and Sweden are quite similar. Sweden is also commonly used as an example of successfully promoting small business development in Europe.

entrepreneurship.[19] Additional details on the data as well as institutional details on the Swedish labor market can be found in Section B of the Appendix.

3 Self-Employment over the Life Cycle

This section provides empirical evidence on the interaction between career dynamics and self-employment decisions. First, I document the variety of ways self-employment fits into life-cycle employment profiles and build a data-driven taxonomy of self-employment behaviors. Second, I use the taxonomy to document the large differences between self-employment behaviors in terms of both background characteristics and initial capital. Third, I develop an intuitive two-period model that rationalizes the various behaviors seen in the data.

Constructing and visualizing life-cycle employment profiles: Using Swedish men born in 1970, I construct life-cycle employment profiles consisting of the sequence of states each person visits over time. Using these profiles, I visualize and compare the careers of everyone who is ever self-employed.[20]

Figure 1 shows the trajectories for all Swedish men born in 1970 who ever report being self-employed. In each year, agents can be in paid employment (PE), unincorporated self-employment (SE-U), incorporated self-employment (SE-I), school (SCH), or non-employment (NE). The top panel shows all the life-cycle employment profile involving any self-employment. Each horizontal line visualizes the states visited by a particular individual between ages 20 and 43, where the colors correspond to which state the individual is in at that age. The panel demonstrates that there is substantial heterogeneity in these profiles. The bottom panel shows the distribution of states within each year. The panel shows that the proportion of the population in self-employment is growing over time while schooling and (to a lesser degree) non-employment decrease over time.

Figure 2 displays three examples of the life-cycle employment profiles shown in Figure 1. Each example shows which state an individual occupied each year from age 20 to age 43. In the

[19]See (Goetz et al., 2015), who argue that such linked data is crucial for understanding small business formation and growth.

[20]This is related to previous literature on life-course analysis. See, for example, Elder and Rockwell (1979) and the subsequent literature reviewed in Jr et al. (2003). Abbott (1983) similarly proposes comparing life cycle profiles across individuals, which is commonly called trajectory analysis.

first example, the individual spent the first 14 years of their career in paid employment (PE) with intermittent non-employment (NE). At age 35, the individual enters into incorporated self-employment (SE-I) and remains there. In the second example, the individual spends his/her entire career in paid employment except for one year spent in incorporated self-employment at age 38. In the final example, the individual spends the majority of the time no-employed but spends a few years in paid employment and two years in unincorporated self-employment (SE). The three examples demonstrate the variety of ways self-employment can fit into a person's life-cycle employment profile but provides no guidance on how common each spell may be.

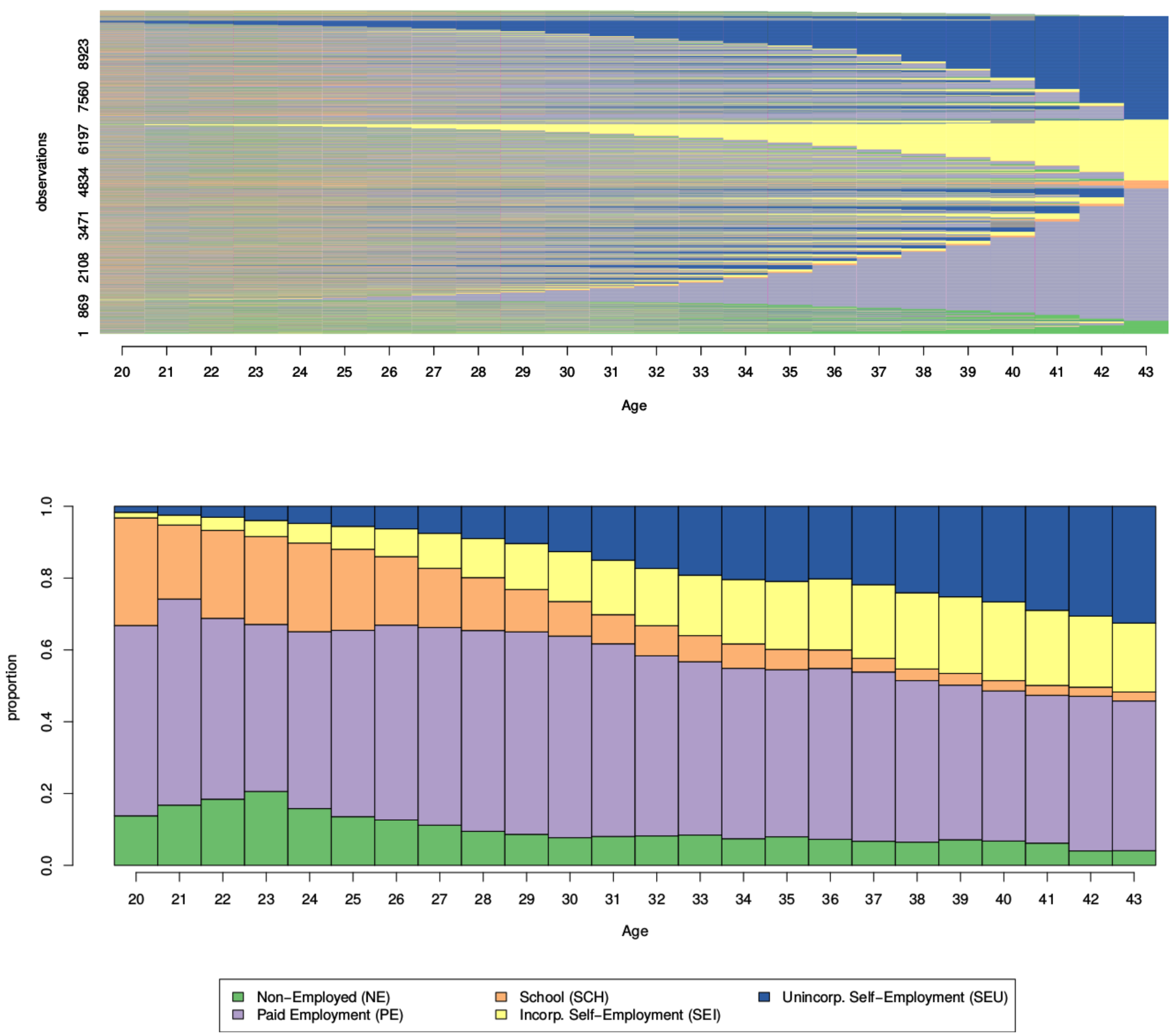
Clustering life-cycle employment profiles: Abbott and Forrest (1986) and the subsequent literature have developed methods for constructing dissimilarities between sequences which allow for similar trajectories to be clustered.[21] The most common method for estimating distances between discrete time series is "Optimal Matching" (OM), which provides a measure of "edit distance" between each pair of trajectories.[22] Specifically, given two trajectories, OM calculates the least costly way to convert one string into another using (1) substitutions, (2) deletions, and (3) insertions, where each of the three operations has an associated cost. The edit distance is then the minimum cost needed to convert one string into another. Optimal Matching can also be thought of as a specific case of an "optimal transport" problem or "Monge-Kantorovich" transport problem (Villani, 2003; Galichon, 2016).

To provide a short example of how the algorithm works, consider the following two short example profiles: $T_1 = PE - NE - PE - PE - SE$ and $T_2 = NE - PE - PE - SE - SE$. In addition, assume deletion costs, addition costs, and substitution costs are all set to unity. T_1 can be transformed into T_2 by (a) switching period 1 to NE, switching period 2 to PE, and switching period 4 to SE; or (b) by deleting period 1 and inserting "SE" to the end of period 5. The first approach has a cost of 3 while the second has a cost of 2, so the optimal matching distance between T_1 and T_2 is 2. When insertion and deletions are relatively more expensive than substitutions, the algorithm places more importance on timing and less importance on

[21] See Studer and Ritschard (2016) for an overview of these methods. See Gabadinho et al. (2009) and Gabadinho et al. (2011) for details on their use and implementation.

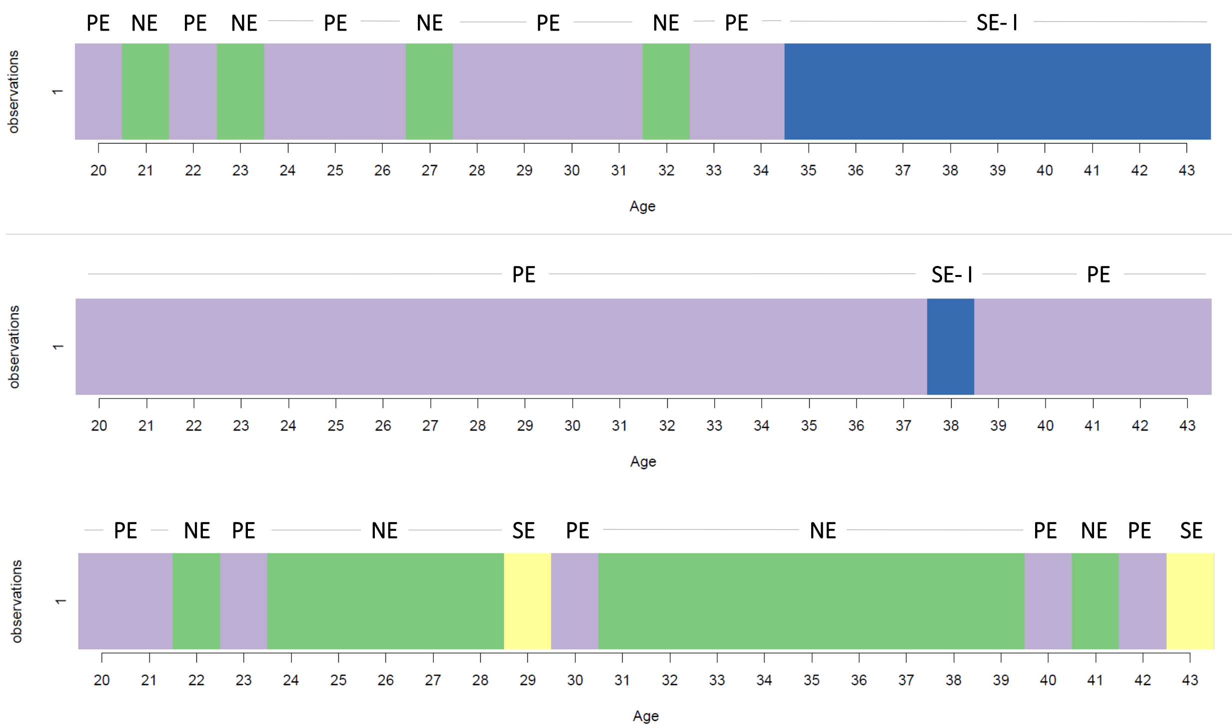
[22] Edit distance is a way of measuring the dissimilarity between two strings based on counting the minimum number of operations needed to transform one string into the other. Which operations are used can vary, but the most common (and ones used here) are substitution, deletion, and insertion.

Figure 1: Life-Cycle Career Profiles Involving Self-Employment (1970 birth cohort)



Notes: Figure shows life trajectories of all Swedish males born in 1970 who are ever self-employed between 1990 and 2013. The top panel shows the life trajectories of all 10,303 individuals born in 1970 who are ever self-employed from age 20 to 43. The bottom panel shows the proportion of observations in each state in each year. NE represents non-employment, SE represents self-employment, and PE represents paid employment. In the top panel, observations have been sorted vertically starting at the end of the panel and working backwards in order to roughly group similar trajectories together. Both panels show only those born in 1970 who ever report being self-employed.

Figure 2: Three Example Career Trajectories Involving Self-Employment.



Notes: Figure shows three example life cycle profiles involving self-employment for the ages 20 to 43. "NE" stands for non-employment, "PE" stands for paid employment, "SE" stands for unincorporated self-employment, and "SE-I" stands for incorporated self-employment.

the sequencing of events.[23]

Given a matrix of the pairwise distances between trajectories, we can cluster using standard hierarchical clustering algorithms. Using Ward's method, I cluster the trajectories shown in Figure 1 into seven distinct groups.[24]

Figure 3 displays the full set of employment profiles grouped into seven unique clusters. The top row shows life cycles that involve late-onset incorporated self-employment or late-onset unincorporated self-employment. The second row shows life cycles where agents spend most of their lives in either incorporated self-employment or unincorporated self-employment. The first sub-figure on the third row shows individuals who spend most of their career working in paid employment but also have a short self-employment spell. The second sub-figure on the second row shows individuals who spend the majority of their careers not working but also occasionally work in paid employment or unincorporated self-employment. The last sub-figure shows individuals who remain in education for much of their early careers and then have late-onset self-employment spells.

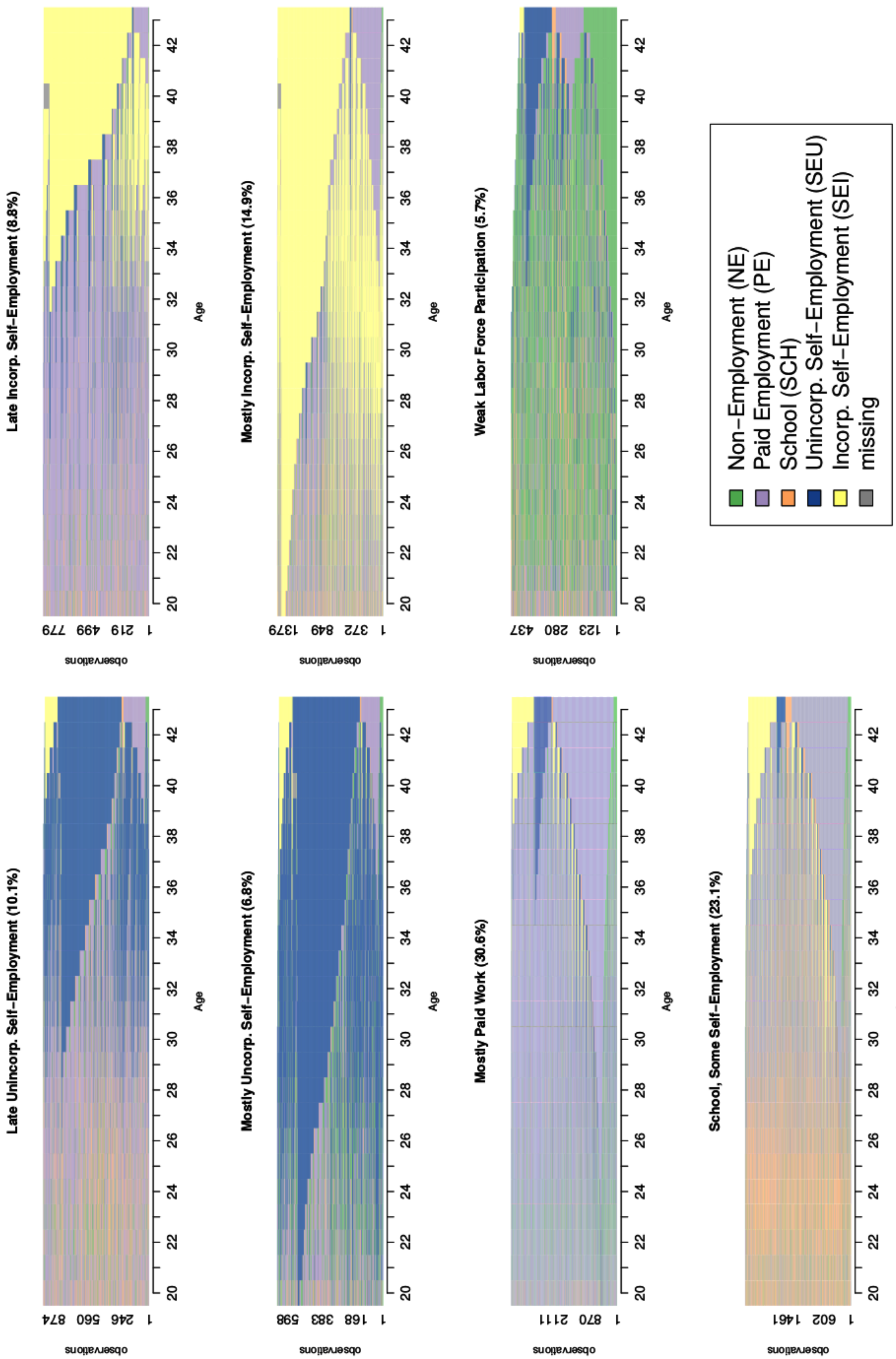
Differences among self-employment clusters: The clusters were created using only information on individuals' life-cycle employment profiles. However, the different clusters vary along other key dimensions. Specifically, (a) the clusters have different earnings trajectories over their careers, (b) members of each cluster come from different backgrounds, and (c) members of each cluster make different decisions upon entering self-employment.

Figure 4 shows that the different associated profile clusters also have different earnings trajectories. The figure shows the average earnings by age for each group as well as for those who never self-employ. Those with weak labor force participation have no wage growth while those who enter unincorporated self-employment early or late earn notably less than those who never self-employ. On average, both groups involving incorporation earn notably more than those who never self-employ. The figure highlights how the different self-employment behaviors

[23]See Section D of the Appendix for details on how dissimilarities are calculated and how costs are chosen.

[24]Ward's method is an agglomerative method for building hierarchical clusters. Starting with each observation as its own cluster, Ward's method identifies which two clusters can be combined with the smallest increase in variance within the cluster. This is repeated recursively, updating the matrix of distances between the new set of clusters using the Lance-Williams Algorithm. Hierarchical clustering produces a tree of nested clusters. The choice of seven clusters was made based on (1) the improvement in predicted lifetime earnings from adding additional clusters and (2) the improvement in fit within the clusters from adding additional clusters. Both flatten out at seven, suggesting seven clusters summarize the data well.

Figure 3: Clusters of Life Cycles Involving Self-Employment (1970 birth cohort)



Notes: Figure shows life employment profiles of all Swedish males born in 1970 who are ever self-employed between 1990 and 2013.

are related to very different earnings trajectories. If the diverse profiles were pooled into a "self-employment" category, their average would be much closer to that of paid employment, but with much larger cross-sectional variance.[25]

Table 1 summarizes the background characteristics of each self-employment cluster. Large differences in abilities exist across clusters. For example, those who start incorporated businesses early in their careers have above average cognitive and non-cognitive ability, those that start unincorporated businesses early in their careers have below average cognitive and non-cognitive abilities, and those with weak labor-force participation have substantially lower abilities. All of the self-employment clusters are more likely to have self-employed parents, but this is notably higher for those who enter self-employment early.[26] The table also documents that there are large differences in the present discounted value of earnings (from ages 18 to 43 discounted at 5%) across the groups.[27] Those who mostly work and those who spend most of their careers in unincorporated self-employment are substantially less likely to have college degrees.

Table 2 provides evidence similar to Table 1 by running binary logistic models for cluster membership for each of the clusters. For comparison, the last two columns show logistic regressions for ever being self-employed and ever being incorporated self-employed. The table shows that cognitive ability is negatively related to self-employment for most clusters except for those who attend school and later self-employ. Similarly, the incorporated self-employed have positive coefficients on non-cognitive ability, yet negative coefficients on cognitive ability. In contrast, if we had pooled all the self-employed workers, we would have conflated the relationship between abilities and background and the various self-employment behaviors. For example, the coefficient on cognitive ability is positive on the "Ever SE" regression despite the coefficient being negative in many of the cluster-specific regressions.

Table 3 documents the self-employment decisions in the first years of self-employment by self-employment group. First, across clusters, there are large differences in the amount of capital employed by businesses in their first year in operation. Those starting incorporated businesses

[25]Figure 11 in the Appendix shows the cross-sectional variation in self-employment is larger than that in paid employment and grows over time. Figure 12 shows the similar figure, but separated out by profile clusters. The variation in earnings across the distinct self-employment behaviors could easily be misinterpreted as increased risk, rather than increased heterogeneity.

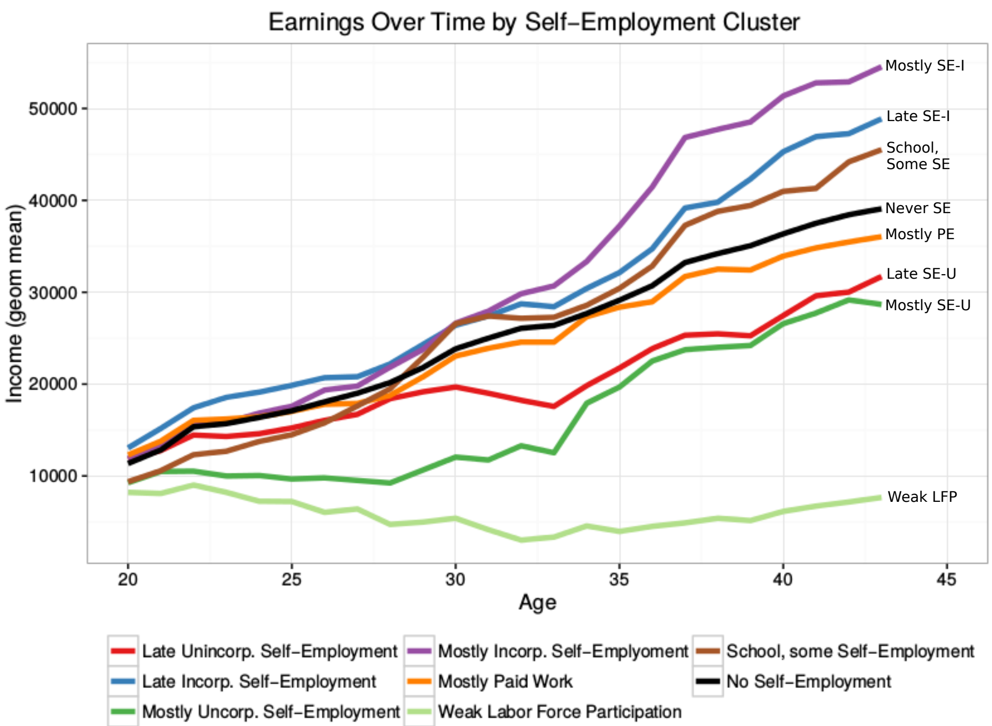
[26]See Lindquist et al. (2015) and Vladasel et al. (2021) for additional evidence on parental and early life influences on entrepreneurship in Sweden.

[27]All monetary amounts in this paper have been converted to 2010 U.S. dollars

employ substantially more capital than the other self-employment clusters which employ minimal levels of capital. Similarly, the self-employed in clusters involving unincorporated self-employment employ few employees. Those who start incorporated businesses later in their career and those who mostly work with a brief self-employment spell are more likely to start businesses in the industry in which they previously worked or in industries in which they have substantial experience. Those who start unincorporated businesses later in their career and those with weak labor-force participation are more likely to enter self-employment from non-employment.

Overall, the self-employment groups differ substantially in their backgrounds and earnings over their lives. Moreover, the clusters differ in terms of what type of businesses they create, how much capital and labor they initially employ in their businesses, and how they enter self-employment. This evidence suggests that the clusters identify meaningfully different groups and that a model aimed at accurately describing self-employment in the economy must be able to justify the heterogeneous self-employment behaviors shown above.

Figure 4: Earnings Trajectories By Self-Employment Cluster



Notes: Figure shows the geometric average after-tax earnings of the various self-employment groups by age. Results are for Swedish men born in 1970, and all results are in 2010 US dollars.

Table 1: Summary of Background Characteristics by Self-Employment Behavior

	Late SE Incorp	Late SE	Mostly SE Incorp	Mostly SE	Mostly Work	Weak LFP	School, some SE	Never Self-Emp
Mean PV Wage Income	405,281	311,424	458,260	266,232	367,063	202,358	402,244	363,092
Mean Wage (age 40)	54,424	33,947	68,636	34,393	43,311	19,233	54,660	44,102
Non-cog Abil	0.17	0.02	0.20	-0.14	0.00	-0.50	0.38	-0.01
Cog Abil	0.04	0.14	0.12	-0.15	-0.10	-0.31	0.71	0.08
Self-Emp Parents	0.53	0.56	0.67	0.66	0.51	0.51	0.52	0.36
Mean Parental Inc	35,419	34,787	37,973	32,176	34,556	33,715	40,065	35,592
Prop Coll Grads	0.04	0.11	0.09	0.01	0.02	0.04	0.32	0.15
N	862	961	1,466	654	3,434	450	2,316	33,538

Notes Table shows summary statistics on ability and background by group. The first seven columns correspond to the self-employment clusters identified above while the eight column presents statistics for those who never self-employ. The sample comprises Swedish men born in 1970. All monetary amounts are in 2010 US dollars. Parental Income is the average income of the parent when the child is between the age of 13 and 18. Cognitive and non-cognitive ability are normalized to have a mean of zero and a standard deviation of one.

Table 2: Predicting SE Behaviors by Ability and Background (binary logistic regressions)

	Late SE Incorp	Late SE	Mostly SE Incorp	Mostly SE	Mostly Work	Weak LFP	School, some SE	Ever SE	Ever Incorp. SE
Non-cog Abil	0.244***	−0.002	0.191***	−0.040	0.128***	−0.349***	0.142***	0.123***	0.251***
	(0.043)	(0.039)	(0.033)	(0.048)	(0.022)	(0.060)	(0.026)	(0.014)	(0.018)
Cog Abil	−0.091**	0.069*	−0.075**	−0.074	−0.112***	−0.130**	0.500***	0.047***	0.014
	(0.043)	(0.041)	(0.033)	(0.049)	(0.023)	(0.061)	(0.029)	(0.015)	(0.019)
Self-Employed Parents	0.502***	0.534***	1.078***	0.995***	0.514***	0.473***	0.379***	0.744***	0.677***
	(0.073)	(0.070)	(0.060)	(0.089)	(0.039)	(0.106)	(0.046)	(0.025)	(0.031)
Log Avg. Parental Wealth	0.125***	0.109***	0.167***	0.146***	−0.023	−0.013	0.069***	0.077***	0.104***
	(0.037)	(0.035)	(0.029)	(0.041)	(0.017)	(0.045)	(0.024)	(0.012)	(0.016)
Log Avg. Parental Income	−0.014	−0.376***	0.545***	−0.813***	−0.013	−0.424***	0.280***	0.015	0.396***
	(0.122)	(0.095)	(0.095)	(0.119)	(0.065)	(0.135)	(0.078)	(0.040)	(0.053)
College Mom	−0.075	0.272***	−0.078	−0.153	−0.275***	0.374***	0.350***	0.083***	−0.005
	(0.098)	(0.084)	(0.072)	(0.122)	(0.056)	(0.136)	(0.051)	(0.031)	(0.039)
Constant	−5.228***	−1.310	−11.458***	2.536**	−1.901***	0.191	−7.376***	−2.289***	−7.501***
	(1.261)	(1.011)	(0.980)	(1.259)	(0.667)	(1.420)	(0.804)	(0.419)	(0.551)
Obs.	43,780	43,780	43,780	43,780	43,780	43,780	43,780	43,780	43,780

Note: Tables shows estimates form binary logistic regressions on selecting into the various types of self-employment on background and ability characteristics. Controls for education are also included but not displayed. The types of self-employment correspond to the life cycle profile clusters estimated above. The last two columns show the estimates from a logit on ever being self-employed and a logit on ever being incorporated. The sample is all Swedish men born in 1970.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 3: Characteristics upon Entering Self-Employment by Self-Employment Cluster

	Late SE Incorp	Late SE	Mostly SE Incorp	Mostly SE	Mostly Work	Weak LFP	School, some SE
Med Fixed Assets	\$28,993	\$3,788	\$44,133	\$6,552	\$6,974	\$2,410	\$4,899
Med Revenue	\$169,977	\$32,389	\$196,348	\$30,584	\$57,311	\$16,775	\$86,291
Avg Employees	4.1	0.2	7.2	0.1	2.0	0.0	4.8
Med Employees	2	0	3	0	0	0	0
Same Indust as last yr	0.67	0.31	0.56	0.27	0.45	0.11	0.52
Norm. Career Exp	0.59	0.46	0.62	0.46	0.45	0.31	0.63
Not working last yr	0.08	0.29	0.20	0.47	0.17	0.71	0.18

Notes: Table shows average business characteristics of the self-employed during their first year in business by self-employment group. “Same Indust as last yr” is the proportion that enter self-employment in the same industry they were in the previous year. “Norm. Career Exp” is the proportion of prior experience acquired in the same industry as the one in which the individual enters into self-employment.

3.1 Interpreting the descriptive evidence

This section presents an intuitive two-period model of self-employment decisions to interpret and rationalize the patterns reported in the previous sub-section. Using the two-period model, I interpret the empirical relationships between individuals' career dynamics and self-employment decisions. In particular, I study the relationship between initial investment and business survival.

Earnings in paid employment and self-employment: Assume the per-period income of choosing paid employment is given by

$$w(A, e_{pe,t}, e_{se,t}) + \epsilon_{pe,t} \tag{1}$$

where w is the expected earnings in paid employment which depends on ability A, accumulated experience in paid employment $e_{pe,t}$, and accumulated experience in self-employment $e_{se,t}$. $e_{pe,t}$ and $e_{se,t}$ are both zero in the first period. $\epsilon_{pe,t}$ is the idiosyncratic shock to paid employment earnings. Similarly, per-period profit in self-employment is given by:

$$\theta(A, e_{pe,t}, e_{se,t})K_t^{\alpha} - rK_t + \epsilon_{se,t} \tag{2}$$

where θ is the expected contribution of the founder's abilities and experiences to the productivity of the firm, K is the amount of capital employed by the firm, and r is the rental rate of capital.[28] I assume that capital is rented on the spot market each period at the common interest rate r but that there is an upward adjustment cost (that scales with K_t). Thus, if an agent chooses to enter self-employment in period t with K_t capital, he must pay a start-up cost of γK_1.[29] Adjustment costs capture the idea that there are costs of entry that scale with size, such as the difficulty of screening and hiring a large number of employees, or working out a large number of rental contracts for capital. As adjustment costs may capture psychic or effort costs that do not directly affect observed earnings, I include them as part of the utility function, but not the

[28] Hamilton (2000) and Hurst and Pugsley (2011) argue that there are large non-pecuniary benefits to self-employment. While excluded in this simple example, the full model allows for heterogeneous non-pecuniary benefits from self-employment.

[29] Given this set-up, there are two reasons individuals would differ in their level of capital investment. First, differences in skills or experiences will result in different K_t^*, but, in period 1, differences in the probability of remaining self-employed in the future will also result in different choices of K_1^*.

income and profit functions shown above.

Value functions: In period 1, the agent's expected discounted utility from choosing paid employment is given by:

$$v_1(pe, K_1, A, 0, 0) = w(A, 0, 0) + \epsilon_{pe,1} + \beta E[V(A, K_1, 1, 0)] \tag{3}$$

where $K_1 = 0$ in paid employment. If the agent chooses to self-employ in period two, the optimal level of capital will be: $K_2^* = \left(\frac{\alpha\theta(A,1,0)}{r+\gamma}\right)^{\frac{1}{1-\alpha}}$. Alternatively, the agent's expected discounted utility in self-employment is given by:

$$v_1(se, K_1, A, 0, 0) = \theta(A, 0, 0)K_1^{*\alpha} - (r + \gamma)K_1^* + \epsilon_{se,1} + \beta E\left[V(A, K_1, 0, 1)\right] \tag{4}$$

where $K_2^* = \left(\frac{\alpha\theta(A,0,1)}{r+\gamma}\right)^{\frac{1}{1-\alpha}}$ will be the optimal level of capital assuming θ is increasing in self-employment experience (i.e. $K_2^* > K_1^*$).

Choice of capital: K_1^* is optimally chosen based on the first order condition:

$$\alpha K_1^{\alpha-1}\theta(A, 0, 0) = r + \gamma + \frac{\partial}{\partial K_1^*} E\left[V(A, K_1, 0, 1)\right]. \tag{5}$$

To simplify the example further, assume that $\epsilon_{se,t}$ and $\epsilon_{pe,t}$ both have type-1 extreme value distributions, the first order condition with respect to K_1 can then be written as

$$\alpha K_1^{\alpha-1}\theta(A, 0, 0) = r + \left(1 - \frac{\beta}{1 + exp\left(\bar{v}_2(pe, A, 0, 1) - \bar{v}_2(se, A, 0, 1)\right)}\right)\gamma. \tag{6}$$

Thus, the optimal choice of K_1 depends on the expected earnings differential in utility between employment and self-employment in the next period ($\bar{v}_2(pe, A, 0, 1) - \bar{v}_2(se, A, 0, 1)$). When the utility gap between paid employment and self-employment is large and positive, optimal K_1 is the solution to $\alpha K_1^{\alpha-1}\theta(A, 0, 0) = r + \gamma$. When the utility gap between paid employment and self-employment is large and negative, the optimal K_1 is the solution to $\alpha K_1^{\alpha-1}\theta(A, 0, 0) = r + (1 - \beta)\gamma$. Thus, the effective interest rate faced by a new business depends on the probability that the individual will remain self-employed in the future. Given the type-1 extreme value assumption, the first-order condition can be rewritten as:

$$\alpha K_1^{\alpha-1}\theta(A,0,0) = r + (1 - \beta Pr(j_2 = se|A,0,1))\,\gamma, \tag{7}$$

where $j_2 \in \{pe, se\}$ is the agent's choice in the second period, and the right-hand side is the shadow price of capital.

In the two-period model, a number of factors influence why agents choose a particular career in the first period and how much they choose to invest if self-employed. Agents may forgo earnings in the first period in order to acquire additional human capital in a specific career. Alternatively, agents may choose to self-employ in the first period in order to avoid negative labor market shocks in paid employment, knowing that they will most likely return to paid employment in the future. Which behaviors are observed depends on the importance and relative transferability of human capital and the relative returns to ability. The amount of capital used in self-employment in the first period will depend on the individual's absolute productivity as well as their relative productivity in self-employment compared to paid employment. Individuals who expect to have higher earnings in paid employment in the future will initially start smaller businesses if they choose to be self-employed.

Testable assumptions: Within this simple framework a number of testable assumptions emerge which are consistent with the data.

1. Brief spells of self-employment are more common in industries that are less capital intensive. Figure 5 plots the relationship between average initial capital investment and average years survived by three-digit industry code. Each dot represents a particular three-digit industry code and is scaled by the number of self-employed individuals in that industry. The blue line shows the OLS linear relationship between average initial capital investment and average years survived. The relationship between average survival and average initial investment across industries is strong and positive, with a correlation of 0.70. This provides evidence that brief spells of self-employment tend to be more common in less capital intensive industries.

2. Conditional on industry, agents who only enter self-employment briefly employ less capital upon entering self-employment. While there is a relationship between self-employment spell length and tenure across industries, it is useful to test this relationship at the individual level. As the model suggests, conditional on the choice of industry, those who enter self-employment briefly should choose to employ less capital. Figure 7 plots the relationship between years survived and total assets employed by the business during its first year in operation. The figure shows that there is a strong relationship between survival and initial assets. Table 4 documents that similar trends hold within most industries, suggesting that agents are forward looking and initially start larger businesses based on their expectations of remaining self-employed in the future.

3. Some self-employment spells may be caused by labor market shocks. Such spells should be short and involve little investment. The top panel of Figure 6 plots the average earnings by self-employment tenure for those who remain self-employed for one year, three years, five years, and seven years. The figure shows that while surviving firms grow, they also earn more during their first year in business. The bottom panel fills in the top panel with the average earnings of each group before and after the self-employment spells shown, using triangles to indicate periods outside of the self-employment spell. For those who only enter self-employment for one year, there is a notable dip in earnings leading into self-employment followed by a rapid recovery. In contrast, those that remain self-employed seven years have increasing earnings upon entering self-employment but a large drop in earnings seven years later upon exit. Of the four groups, it is those who only self-employ for one year who have the highest earnings three years prior and within two years after starting their business. This suggests that this group may have entered self-employment only due to labor market shocks and that their high earnings in paid employment quickly induced them to leave self-employment.

4. Brief spells of self-employment should be more common in industries where self-employment experience and paid employment experience are close substitutes. Table 5 shows Mincer-regressions for paid employees by major industry groups. The table breaks experience down into industry-specific experience in paid employment, industry-specific experience in self-employment, and other work experience. The regressions additionally control for ability, background, and education, though these coefficients are not displayed. For the industries with high rates of self-employment, such as construction, financial and business services, and trade and communication, the returns on a year of self-employment experience are almost equal to the returns to a year of paid employment experience.

Overall, the empirical evidence suggests that the career dynamics of individuals have important implications for the businesses they choose to create. The data supports the simple model which demonstrates that many low-profit and short-lived businesses may be a natural consequence of individuals smoothing over labor market shocks while continuing to acquire human experience within their industry. Individuals across the distribution of abilities face different opportunity costs when choosing to self-employ based on the relative value of their skills and abilities in paid employment. This explains why we find both low- and high-skill individuals entering self-employment. This is also consistent with the analysis of Wennberg et al. (2010) which highlights that many founders exit well-performing businesses due to better opportunities in paid employment.

The evidence presented here also agrees with the evidence from Section 3 where, for at least two of the documented self-employment behaviors, the individuals appear to have little intention of remaining self-employed in the future and make different self-employment decisions accordingly. The simple model and supporting evidence also provides an alternative explanation for one of the two motivating facts from Hopenhayn (1992), who observed that "entry and exit rates are highly correlated across industries and most of their variation is accounted for by these industry effects." This model provides a number of reasons why self-employment spells may be shorter in some industries than others due to capital intensity and the transferability of experience between self-employment and employment in that industry.

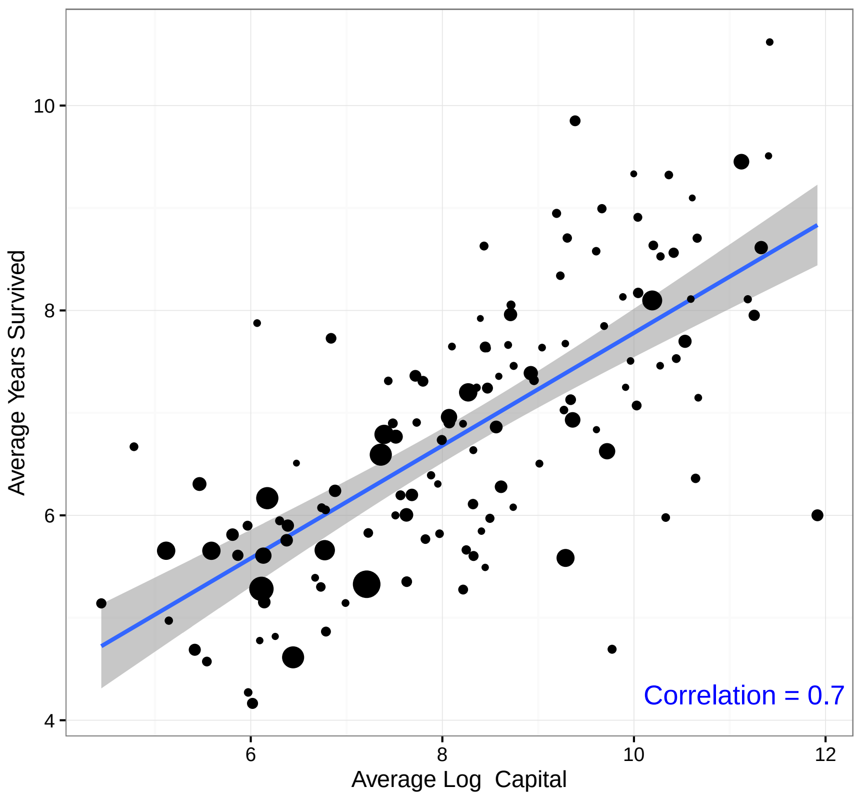
The two-period model provides alternative explanations for two other empirical regularities seen in the data. First, a model with capital adjustment costs provides an alternative explanation of the "up-or-out" patterns for new small businesses where new low-profit businesses tend to be short-lived.[30] In particular, the relationship between survival and earnings is commonly explained by models where agents learn about their entrepreneurial ability over time. [31] Second,

[30]See Haltiwanger (2012) for a recent review on firm dynamics which documents that many young small businesses fail but that some young businesses that do not fail grow rapidly.

[31]See, for example, Jovanovic (1982), Dillon and Stanton (2016), (Manso, 2016), and Hincapi (2020).

the life-cycle model explains why short-lived businesses may appear to "under-invest" in their businesses, evidence that has been previously interpreted as credit constraints affecting the intensive margin of new businesses. While my model does not provide a direct test against the previous explanations for these empirical regularities, it provides an alternative framework capable of explaining both.

Figure 5: The Relationship Between Average Years Survived and Average Initial Capital by Industry (3 digit)

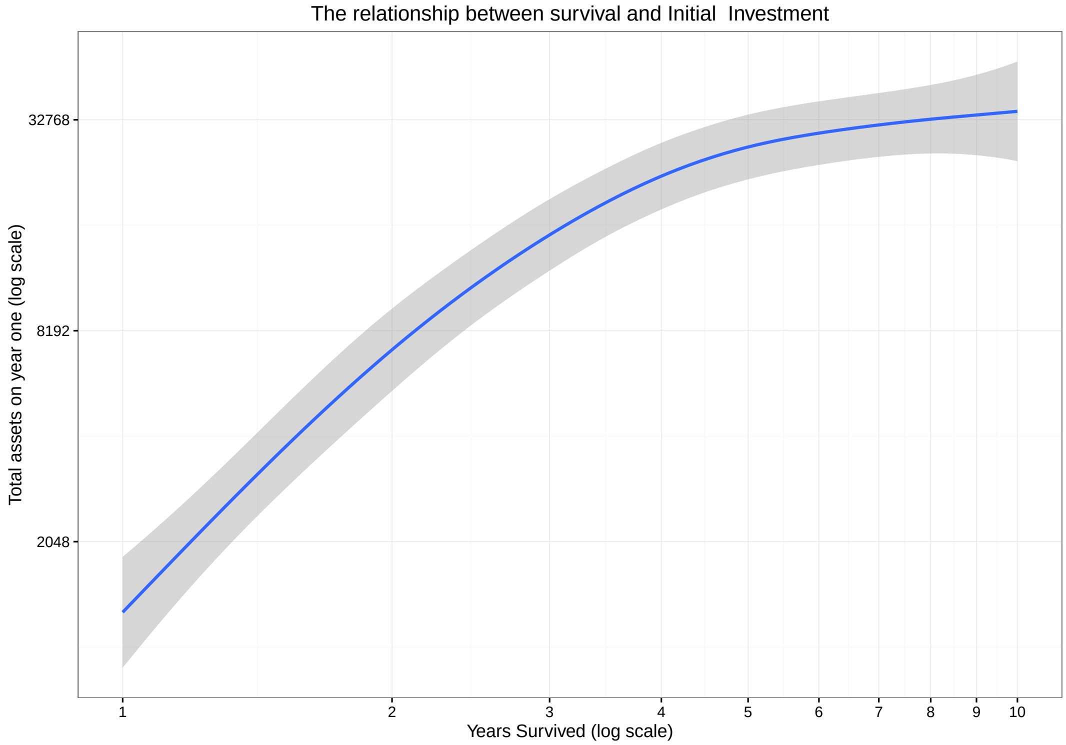


Notes: Figure shows a scatterplot of the average log initial capital employed by new firms and the average years survived by industry (3 digit). Only industries in which I observe 50 or more new small businesses enter are shown. Dots are scaled by the number of new small businesses observed in that industry. The blue line shows the linear trend

Figure 6: Earnings Dynamics for the Newly Self-Employed who Stay 1, 3, 5, and 7 years

Notes: Figure displayed log total work-related compensation for new businesses based on the number of years they survive. The data consists of new self-employment spells started by Swedish men born between 1965 and 1975 who started their businesses between 1990 and 2005. The numbers represent the number of years the firm survived. Log income is in 2010 U.S. dollars.

Figure 7: Relationship between Initial Investment and Firm Survival



Notes: Figure shows the relationship between initial capital investment and the number of years a firm survives. The fitted line is estimated for individuals who entered self-employment between the ages of 30 and 36 and who entered no later than 2002. The line is fitted using a LOESS regression and the grey area shows the standard error of the estimate.

Table 4: Initial Capital Investment by Years Survived by Industry

Years Survived	Agriculture	Manufac. and Mining	Construction	Trade and Commun.	Finance and Bus. Services	Personal Serv.
1	$2,943	$3,771	$2,343	$4,211	$1,908	$1,678
2	$4,696	$4,964	$3,304	$6,475	$3,618	$2,800
3	$5,760	$4,389	$3,240	$6,288	$4,454	$2,657
4	$7,360	$4,811	$3,810	$8,106	$4,149	$2,863
5	$12,674	$6,450	$4,675	$9,506	$6,138	$3,949
6	$12,288	$5,175	$4,190	$9,453	$7,746	$2,257
7	$13,583	$6,125	$4,733	$12,917	$8,267	$2,851
8	$16,174	$8,300	$5,000	$14,886	$8,483	$2,793

Notes: Figure shows the median amount of fixed assets employed by new firms in their first years of business split out by years the firm survived and major industry group. All monetary amounts are expressed in 2010 U.S. dollars.

Table 5: Is Experience in Self-Employment Valued In Paid Employment? Evidence from Earnings Regressions by Industry

	Agriculture	Manuf. and Mining	Energy, Water, and Waste	Construction	Trade and Commun.	Finance and Bus. Services	Edu and Research	Health and Social Care	Personal and Cultural Serv.	Public Admin
Career Exp (PE)	0.028***	0.025***	0.017***	0.031***	0.032***	0.041***	0.022***	0.033***	0.037***	0.026***
	(0.003)	(0.0004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Career Exp^2 (PE)	−0.0004***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.0001)	(0.00002)	(0.0001)	(0.00003)	(0.00003)	(0.00005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Career Exp (SE)	0.019***	0.028***	−0.057***	0.030***	0.024***	0.046***	−0.030***	−0.027***	0.008***	−0.012**
	(0.003)	(0.001)	(0.007)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)	(0.003)	(0.005)
Other Work Exp	0.017***	0.008***	0.003***	0.011***	0.016***	0.008***	0.009***	0.003***	0.015***	0.002***
	(0.002)	(0.0003)	(0.001)	(0.0004)	(0.0003)	(0.0003)	(0.0005)	(0.001)	(0.001)	(0.001)
Obs.	7,642	194,169	7,204	72,398	138,880	104,026	37,632	26,463	25,910	23,502
R^2	0.462	0.617	0.604	0.566	0.545	0.474	0.366	0.603	0.450	0.591
Adj. R^2	0.460	0.617	0.602	0.565	0.545	0.474	0.365	0.602	0.449	0.591

Notes: Tables shows earnings regressions for Swedish men in paid employment between the ages of 20 and 45 who have completed schooling. Regressions are estimated separately for each major industry. "Career Exp (PE)" is the number of years worked in the industry as a paid employee, "Career Exp (SE)" is the number of years the number of years self-employed in the industry, and "Other Work Exp" is all additional years of work experience. Regressions control for cognitive ability, non-cognitive ability, education, major, parents' income, if the parents were ever self-employed, and age dummies. Units are in (log) total after-tax compensation in 2010 U.S. dollars. *p<0.1; **p<0.05; ***p<0.01

4 A Structural Model of Self-Employment Decisions

The two-period model presented in the previous section provides a framework for interpreting the empirical evidence but does not capture the rich patterns of self-employment over the life-cycle previously documented. It also does not consider multiple sectors and does not incorporate the non-pecuniary costs and benefits of self-employment This section lays out a structural model of human capital accumulation that nests the choice to self-employ within a larger life cycle model of career choice. Extending the two-period model, the structural model includes multiple sectors with sector-specific experience, non-pecuniary benefits, and transition costs incurred for switching between sectors. The structural model allows me to quantify the importance of abilities and career histories on self-employment decisions and to estimate policy counterfactuals.

In period t, each person observes a vector of state variables Ω_t and chooses their action $\boldsymbol{a}_t \in A$ in order to maximize their expected discounted lifetime utility. Actions $a_t = \{d_t, K_t\}$ consist of decisions of which industry and organizational form to work in and how much capital to employ in their business if self-employed. Building on a common framework of dynamic discrete choice as laid out in Rust (1994) and Aguirregabiria and Mira (2010), people choose $\boldsymbol{a}_t$ to maximize

$$E\left[\sum_{\tau=t+1}^{T} \beta^{\tau-t} U_t(a_\tau, \Omega_\tau) | a_t, \Omega_t\right] \tag{8}$$

where, by Bellman's principle of optimality, we can express the expected discounted flow of utility recursively as

$$V_t(\Omega_t) = \max_{a_t \in A} \left[U_t(a_t, \Omega_t) + \beta \int V_{t+1}(\Omega_{t+1}) dF(\Omega_{t+1}|a_t, \Omega_t)\right], \tag{9}$$

and the choice-specific value function is given by:

$$v_t(a_t, \Omega_t) = U_t(a_t, \Omega_t) + \beta \int V_{t+1}(\Omega_{t+1}) dF(\Omega_{t+1}|a_t, \Omega_t). \tag{10}$$

Using this framework, I invoke a set of assumptions that are common in the dynamic discrete choice (DDC) literature. See Section E of the appendix for details.

4.1 Market environment, constraints, and parameterizations

Using the dynamic discrete choice framework above, I build a dynamic model of career choice where the within-period model is similar to static models considered in the literature such as Evans and Jovanovic (1989) and Hamilton et al. (2015). To begin, I lay out the flow-utility and earnings equations. Then I develop the Bellman equations associated with the dynamic model.

Period-specific utility

Building on the simple model in Section 3.1, utility in period t is given assumed to depend on consumption C_t and non-pecuniary benefits H_t:

$$\underbrace{U_t}_{\text{flow utility}} = \underbrace{C_t}_{\text{consumption}} + \underbrace{H_t}_{\text{non-penuniary}} \tag{11}$$

where I assume consumption and non-pecuniary benefits are separable. In the model, I assume agents live hand to mouth and that the utility of consumption is proportional to total after-tax income. H_t will capture the non-pecuniary benefits associated with particular career choices, which I allow to vary based on the observable characteristics and latent type of the individual. H_t will also capture non-pecuniary switching and adjustment costs associated with career changes.

Agent chooses actions $\mathbf{a}_t = \{d_t, K_t\}$ where $d_t \in \boldsymbol{D} = \{0, ..., J\}$ represents choice of industry and organizational form at time t, and K_t is the amount of capital to employ in the firm (0 if not self-employed). There are N_k possible industries in which a person can work $J = 3N_k + 1$ possible choices as agents can choose to be in paid employment, unincorporated self-employment, or incorporated self-employment within each industry. For convenience let choices $\{1, ...N_k\}$ be paid employment, $\{N_k + 1, ..., 3N_k\}$ be unincorporated self-employment, $\{N_k + 1, ..., 3N_k\}$ be incorporated self-employment, and 0 be non-employment. Furthermore, let $h_t \in \{ne, pe, sei, seu\}$ represent which form of labor force participation was chosen at time t. $\mathbf{Z}_t$ are the observable state variables, while $\mathbf{X}_t \in \mathbf{Z}_t$ are observable state variables that affect earnings, $\mathbf{W}_t \in \mathbf{Z}_t$ are observable state variables that may affect an agent's ability to rent capital, and η is a persistent unobserved multinomial state variable. Consumption at time t will then be given by:

$$C_t = \rho Y(\boldsymbol{a}_t, \boldsymbol{X}_t, \boldsymbol{W}_t, \eta, \xi(\boldsymbol{a}_t)) + \rho\tau_t \tag{12}$$

where $\xi(d_t)$ is a idiosyncratic productivity shock and $Y(\boldsymbol{a}_t, \boldsymbol{X}_t, \boldsymbol{W}_t, \eta, \xi(\boldsymbol{a}_t))$ is after-tax earings which will depend on the choice of d_t and, if self-employed, the choice of capital K_t. τ_t captures government subsidies or other transfers. Given the assumption that agents live hand to mouth, Equation 12 also represents the period-specific budget constraint.

Earnings and profit

If in paid employment ($h_t = pe$), earnings at time t are given by

$$Y_t(\boldsymbol{a}_t, \boldsymbol{X}_t, \boldsymbol{W}_t, \eta, \boldsymbol{\xi_t}) = w_t(d_t, \boldsymbol{X}_t, \eta)\xi(d_t),$$

where $w_t(d_t, \boldsymbol{X}_t, \eta)$ is the deterministic portion of earnings in paid employment and $\xi(d_t)$ is a career-specific idiosyncratic shock. If agents are self-employed ($h_t \in \{sei, seu\}$), profits at time t are given by:

$$Y_t(\boldsymbol{a}_t, \boldsymbol{X}_t, \boldsymbol{W}_t, \eta, \boldsymbol{\xi_t}) = \theta_t(d_t, \boldsymbol{X}_t, \eta)K^{\alpha(d_t)}\xi(d_t) - r(\boldsymbol{W}_t)K_t$$

where $\theta_t(d_t, \boldsymbol{X}_t, \eta)$ is the deterministic portion of firm productivity contributed by the founder and $\xi(d_t)$ is a career specific idiosyncratic shock. As $\theta_t(d_t, \boldsymbol{X}_t, \eta)$ depends on d_t, agent characteristics such as cognitive and non-cognitive ability may differ in value across industries.

Similar to the simple model in Section 3.1, I assume agents rent capital on the spot market. Again, I allow for effort-costs associated with the upward adjustment of capital which I discuss more below. Capital productivity is industry specific. Rental rates depend on observable W_t such as parental wealth, which may affect the rate at which individuals can rent capital. Overall, the within-period earnings and profits are similar to Evans and Jovanovic (1989), but generalized to allow productivity in self-employment and paid-employment to depend on cognitive and non-cognitive skills, career-specific experience, and background characteristics.

Non-pecuniary benefits

Non-pecuniary benefits associated with action $\boldsymbol{a}_t$ are given by:

$$H_t = \underbrace{h_t\left(d_t, \boldsymbol{Z}_t, \eta\right)}_{\text{Non-pecuniary}} - \underbrace{\mathbb{1}\{K_t > K_{t-1}\}\gamma(K_t - K_{t-1})}_{\text{Adjustment Costs}} - \underbrace{\lambda_t\left(d_t, d_{t-1}\right)}_{\text{Switching Costs}} + \underbrace{\epsilon(d_t)}_{\text{pref. shocks}}$$

where $h_t\left(d_t, \boldsymbol{Z}_t\right)$ captures the non-pecuniary benefits of a particular career. This is allowed to depend on agent's unobserved type η as well as observable characteristics of the individual such as cognitive and non-cognitive skills which may affect how much agents enjoy working in a particular industry. $h_t\left(d_t, \boldsymbol{Z}_t, \eta\right)$ may also capture the disutility associated with long-hours or earnings uncertainty associated with some choices of employment, though these aspects will not separately identified.

γ captures the notion of upward adjustment costs to growing one's business. As in the simple model, γ captures the idea that there are effort-costs associated with finding capital to employ or otherwise increasing the size of one's business. These effort costs cannot be recovered upon exiting self-employment and are only paid for capital above and beyond the level of capital used the previous period. The adjustment costs make the choice of capital forward-looking, and will lead to different levels of capital investment between agents who may be equally productive in self-employment depending on their relative productivity in paid employment and the likelihood they exit in the future.[32]

$\lambda_t\left(d_t, d_{t-1}\right)$ captures costs of switching from career d_{t-1} at time $t-1$ to career d_t at time t. While career-specific experience affects the opportunity costs of switching careers, transition costs capture preferences. I assume these costs are incurred each time agents change careers, though the switching costs differ between pairs of careers and are allowed to vary with potential experience.

Timing of shocks and decisions

The model has three unobserved state variables. η is a persistent state variable known by the agents that affects earnings and decisions. $\boldsymbol{\epsilon_t}$ are idiosyncratic taste shocks that are unobserved by the econometrician but affect the career decisions of the agents. $\boldsymbol{\xi_t}$ are idiosyncratic

[32] While the model assumes upward adjustment costs, the model has similar implications if there are alternatively quadratic adjustment costs or shut-down costs.

productivity shocks that affect the earnings or profits of an individual in a given career.

I assume that agent's choose their actions $\boldsymbol{a}_t = \{d_t, K_t\}$ sequentially. At the beginning of the period agents choose d_t knowing their idiosyncratic taste shocks $\boldsymbol{\epsilon(t)}$, but not yet knowing the idiosyncratic productivity shocks $\boldsymbol{\xi(t)}$. After choosing their career, if self-employed ($h_t = se$), agents choose their level of capital K_t also not yet knowing their idiosyncratic productivity shock $\boldsymbol{\xi(t)}$. Thus, the choice of capital helps identify persistent unobserved heterogeneity η.[33]

Dynamics

The beginning-of-period value function used for making career decisions can be written as:

$$V_t(\boldsymbol{Z}_t, \eta, \boldsymbol{\epsilon}_t, \boldsymbol{\xi}_t) = \max_{d_t \in D} \left\{ u_t(d_t, \boldsymbol{Z}_t, \eta, \epsilon(d_t)) + \beta \int \bar{V}_{t+1}(\boldsymbol{Z}_{t+1}, \eta) dF_t(\boldsymbol{Z}_{t+1} | \boldsymbol{Z}_t, d_t, \eta) \right\}$$

where agents maximize discounted expected utility and u_t is $E_{\xi_t}[U_t | d_t, \boldsymbol{Z_t}, \eta]$ where agents additionally know their optimal choice of K_t if choosing to self-employ. $\bar{V}_{t+1}(\boldsymbol{Z}_{t+1}, \eta)$ is the integrated value function, integrated over $\boldsymbol{\epsilon}$ and $\boldsymbol{\xi}$. Additionally, define $v_t(d_t, \boldsymbol{Z}_t, \eta_t) = u_t(d_t, \boldsymbol{Z}_t, \eta, \epsilon(d_t)) + \beta \int \bar{V}_{t+1}(\boldsymbol{Z}_{t+1}, \eta) dF_t(\boldsymbol{Z}_{t+1} | \boldsymbol{Z}_t, d_t, \eta)$ to be the choice-specific value function.

Within-period capital choice

While making career choice d_t, self-employed agents also choose how much capital K_t to rent on a spot market at price $r(W_t)$ in order to maximize expected utility. If entering self-employment, capital will be chosen to maximize:

$$\begin{aligned} K_t^* = \underset{K_t}{\operatorname{argmax}} \quad & \rho\theta_t(d_t, \boldsymbol{X}_t, \eta) K_t^{\alpha(d_t)} \xi(d_t) - \rho r(\boldsymbol{W}_t) K_t - \mathbb{1}\{K_t > K_{t-1}\} \gamma K_t \\ + \quad & h_t(d_t, \boldsymbol{Z}_t, \eta) + \epsilon(d_t) + \beta \int \bar{V}_{t+1}(\boldsymbol{Z}_{t+1}, \eta) dF_t(\boldsymbol{Z}_{t+1} | \boldsymbol{Z}_t, d_t, \eta). \end{aligned} \tag{13}$$

As in the simple two-period two-state model above, if capital affects the continuation value through adjustment costs, agents would simply choose the profit-maximizing amount of capital

[33] In future work it may be possible to relax the assumption that agents do not know $\boldsymbol{\xi}_t$ when choosing K_t. In that setup, there would be two sources of endogeneity: (a) the latent persistent heterogeneity η that affects career choices and outcomes and (b) idiosyncratic productivity $\xi_t(d_t)$. The model presents two possible candidates for instruments for $\xi_t(d_t)$: (1) variables affecting the interest rate (W_t) and (2) known innovations that raise earnings in other careers in the future (i.e. differences in future expected wage differentials.

accounting for the effort costs associated with starting a larger business (γ). Yet, because K_t affects effort costs in the next period, agents will account for how likely they are to remain self-employed in the same industry the next period when optimally choosing K_t. As in the simple two-period model, assuming $\boldsymbol{\epsilon}$ is distributed Type-1 extreme value, the choice of K_t will depend on the probability of shutting down tomorrow.

When solving period t of the dynamic discrete choice problem, agents will make career decisions knowing they will choose the optimal K_t for the various self-employment decisions. Note that the choice of K_t will depend on θ_t and W_t, but also on variables Z_{t+1} affecting career decisions in period $t+1$.

State transitions

In each period we observed state variables $\boldsymbol{Z}_t = \{\boldsymbol{z}_t, \boldsymbol{A}, s, \boldsymbol{E}_t, \boldsymbol{x}_t, \boldsymbol{W}_t, K_{t-1}\}$, where $\boldsymbol{z}_t$ are variables affecting career choices but not earnings, $\boldsymbol{A}$ is a vector of cognitive and non-cognitive ability as measured at age 18, s is an indicator of final schooling, $\boldsymbol{E}_t$ captures career history and is a vector of career-specific years of experience in each career of length J plus an indicator of career in the previous period ($\boldsymbol{E}_t = \{e_1, ..., e_J, d_{t-1}\}$), $\boldsymbol{x}_t$ are other covariates affecting the earnings of individuals such as economic conditions, and $\boldsymbol{W}_t$ are variables that may affect the rate at which the individual can rent capital, such as parental wealth.

Experience and previous occupation evolve in a deterministic manner based on the agent's career decisions. Specifically, experience is measured as years of experience and evolves as

$$e_{j,t} = e_{j,t-1} + \mathbb{1}\{d_{t-1} = j\},\ \forall j \in J \text{ and } t \in 0, ..., T.$$

Schooling s is chosen at period $t = 0$ and abilities $\boldsymbol{A}$ are fixed prior to entering the labor market. Individuals are assumed to know the evolution of state variables that may influence earnings ($\boldsymbol{x}_t$) and observable characteristics that affect career choice but not earnings ($\boldsymbol{z}_t$).

Education at time $t = 0$

To account for the endogeneity of schooling, which may be an important determinant of self-employment and career decisions, the model assumes agents make an educational decision at time $t = 0$. This choice is modeled as a single multinomial educational choice of $s \in \{S_1, ..., S_{N_s}\}$

which depends on flow utility $f_s(Z_0, \eta)$ and expected future gains.

$$V_s(Z_0) = f_s(\boldsymbol{Z}_0, \eta) + \beta \int \bar{V}_1(\boldsymbol{Z}_1, \eta) dF(\boldsymbol{Z}_1 | \boldsymbol{Z}_0, s, \eta) + \epsilon(s)$$

Schooling also provides an additional information about η.

Table 6: Defining observed state variables and decisions

$\boldsymbol{A}$	is a vector of baseline cognitive and non-cognitive ability.
$\boldsymbol{E}_t$	is a vector of years of experience in each possible choice in D and the career choice made the previous period $\boldsymbol{E}_t = \{e_1, ..., e_J, d_{t-1}\})$.
d_t	takes on the values $\{0, ..., J\}$, where $J = 3 * N_k + 1$, depending on which of the decisions was chosen by the agent at time t. For convenience let choices $\{1, ...N_k\}$ be paid employment, $\{N_k + 1, ..., 2N_k\}$ be unincorporated self-employment, $\{2N_k + 1, ..., 3N_k\}$ be incorporated self-employment, and 0 be non-employment
h_t	indicates if the agent was in incorporated self-employment, unincorporated self-employment, non-employed, or in paid employment in a given period ($h_t \in \{ne, pe, sei, seu\}$).
$\boldsymbol{X}_t$	is the set of variables affecting earnings outcomes that are observed by the econometrician, including ability and career experience: $\boldsymbol{X}_t = \{\boldsymbol{A}, \boldsymbol{E}_t, s, \boldsymbol{x}_t\}$.
$\boldsymbol{W}_t$	is the set of state variables that affect the rate at which a person can rent capital.
η	represents a permanent state variable that is unobserved by the econometrician but observed by the agent.$^{(b)}$
$\boldsymbol{Z}_t$	is the full set of state variables observed by the econometrician that affect career decisions $\boldsymbol{Z}_t = \{\boldsymbol{X}_t, \boldsymbol{W}_t, \boldsymbol{z}_t \, K_{t-1}\}$.
K_t	is the optimal level of capital an individual will use based on the full set of observable state variables $\boldsymbol{Z}_t$, their choice d_t, and their unobserved state variable η, $K_t = K(d_t, \boldsymbol{Z}_t, \eta)$.

Notes: (a) Variables appearing in z_t not appearing in x_t represent instruments for the outcomes. (b) It is possible to relax the assumption that η is constant over time, though constant latent "types" are common ways of accounting for unobserved heterogeneity in structural models. For example, see Keane and Wolpin (1997), Arcidiacono (2004), or Stange (2012) for examples of its use in dynamic discrete choice models. Advances presented in Arcidiacono and Miller (2011) and Arcidiacono and Jones (2003) have provided iterative methods using the EM algorithm to estimate dynamic discrete choice models with latent types making models with persistent but not permanent heterogeneity computationally feasible.

5 Estimation

The model laid out in Section 4 is estimated recursively via maximum likelihood. Choice-specific continuation values are constructed using estimates of flow utilities in future periods and semi-parametric estimates of choice probabilities. The estimation exploits the two-period finite dependence present in the model, allowing continuation values to be constructed using only choice probabilities and flow-utilities from two periods into the future.

Latent heterogeneity $\boldsymbol{\eta}$: The estimation accounts for latent heterogeneity across individuals by allowing agents to be one of four unobserved types. These types are integrated out of the likelihood using a sequential EM algorithm as implemented in Arcidiacono and Jones (2003). The model iterates between updating the posterior probabilities that agents are of a given type η and solving the model using these posterior probabilities.[34] As the EM algorithm occurs as an outer loop, the model must be solved every iteration. For the remainder of this section, I will discuss estimation treating η as known.

Dependence between choices and outcomes: The estimation relies on conditional independence assumptions, or that all the correlations between choices and between outcomes and choices are being driven by state variables $\boldsymbol{Z}$ and unobserved state variable η. These assumptions are similar to the assumptions needed for matching, but matching on observable state variables $\boldsymbol{Z}$ and unobserved persistent state variable η. η captures the persistent heterogeneity that drives otherwise identical people to persistently behave differently across their choices and outcomes over time. While η only takes on a finite set of values, it can approximate flexible forms of multi-dimensional unobserved heterogeneity. In the model, the latent heterogeneity affects earnings and non-pecuniary benefits, allowing the latent types to capture both unobserved differences in productivity or unobserved differences in preferences. Given $\boldsymbol{Z}$ and η, many of the conditional likelihoods become separable, allowing the maximization step within the EM algorithm to be divided into many smaller optimization problems.

[34] See Kasahara and Shimotsu (2009) for a proof that these latent types can be identified under the assumptions made in this analysis. Generally, identification is aided by my assumption that η does not vary over time and through restrictions on how the state space evolves.

Finite dependence: Choice-specific integrated value functions are estimated via simulation, taking into account the two-period finite dependence present in the model as laid out in (Arcidiacono and Miller, 2011). Specifically, agents' decisions affect future state variables through their experience vector $\boldsymbol{E}_t$ (i.e. how many years they have worked in each career) and their previous career choice d_{t-1}. Given this structure, for any two particular career choices at time t, there are future career choices which result in the agent having the same expected state-space at time $t+3$.[35] Using the assumption that ϵ has a type-1 extreme value distribution, it is possible to estimate the difference in continuation values associated with any career choice and non-employment using only flow utilities and choice probabilities from two periods into the future. Section F of the Appendix provides additional details on conditional choice probabilities and finite dependence.

Solving for career choice $d(t)$: Within each period, estimation follows the following three steps. First, I estimate how much capital the individual would choose to employ if self-employed. Given $\boldsymbol{Z}$ and η and the assumption, capital decisions can be estimated as:

$$\log(K(d_t)) = g\left(\boldsymbol{X}_t, \boldsymbol{W}_t, \eta, Pr(d_{t+1} = d_t | d_t)\right) + \nu(d_t) \tag{14}$$

where capital choice depends on variables affecting earnings ($\boldsymbol{X}_t$), variables affecting the rental rate of capital ($\boldsymbol{W}_t$), the probability that the individual will leave self-employment in period $t+1$, and measurement error $\nu(d_t)$. To estimate equation 14, I must first estimate $Pr(d_{t+1} = d_t|d_t)$. In particular, this is a function of the state variables in period $t+1$ including Z_{t+1}. As shown in equation 13, the probability of exit tomorrow will depend on capital decisions today, therefore all determinants of $K(d_t)$ (i.e. $\{\boldsymbol{X}_t, \boldsymbol{W}_t, \eta\}$) are included as determinants of the exit probability. Given the estimated exit probability, I then estimate $\bar{K}(d_t)$ for each agent and each career involving self-employment.

Second, I estimate the earnings equations for each career via maximum likelihood and calculate the expected earnings in each career for each individual.

Third, I construct the choice-specific integrated value function for each choice using condi-

[35] For example, if an agent chose blue-collar paid employment, then non-employment, then blue-collar paid employment, they would have the same expected state space at time $t+3$ as if they had chosen non-employment, then blue-collar paid employment, then blue-collar paid employment.

tional choice probabilities and finite dependence. This requires estimating four counterfactual flow utilities and four choice probabilities for each possible career decision. Given the expected earnings and the choice-specific integrated value function, I estimate the multinomial choice model of careers, which, conditional on η, is a multinomial logit. Using this approach, the model is solved recursively starting with the final time period T.

Table 7 lists the variables used in this analysis and the decisions or outcomes in which each variable enters. A check mark indicates that the variable is a direct determinant of the outcome or choice. For example, experience does not directly affect career choices and does not have a check mark, though it does enter indirectly through its impact on expected earnings.

Table 7: Variables Entering Each Decision or Outcome

Variable	Decision or Outcome				
	Choice	Earnings	Capital	Stay Prob.	Education
Cog Ability	✓	✓	✓	✓	✓
Non-cog Ability	✓	✓	✓	✓	✓
Education	✓	✓	✓	✓	x
Parental Income	✓	✓	✓	✓	✓
Parental Education	x	x	x	x	✓
Parental Wealth	x	x	✓	✓	x
Parents Ever Self-Emp.	✓	✓	✓	✓	✓
Exp.	x	✓	✓	✓	x
Industry Exp.	x	✓	✓	✓	x
Career Exp.	x	✓	✓	✓	x
Last Career	x	✓	✓	✓	x
Region	✓	✓	✓	✓	x
Unemployment Rate	x	✓	✓	✓	x
Expected Earnings	✓	x	x	x	x
Expected Capital	x	✓	x	x	x
Stay Prob.	x	x	✓	x	x
Latent Types	✓	✓	✓	✓	✓

Notes: This table lays out which variables are used in the analysis and which choices and outcomes they enter. "Stay Prob." refers to the estimated probability of staying in self-employment at time $t+1$ conditional on entering self-employment at time t.

6 Empirical Results

This section presents six sets of results from the structural model. The first evaluates the model's ability to generate the rich patterns of self-employment behavior over the life-cycle seen in the data. The second evaluates how changing skills affects the level and types of self-employment behavior over the life-cycle. The third considers a number of counterfactual policies designed to promote self-employment. The fourth evaluates the importance of careers in determining self-employment decisions. In particular, it evaluates the transferability of experience and transition costs between carers. The fifth set of results evaluates the determinants of capital investment. The sixth set of results discusses the role of latent heterogeneity η in the model.

6.1 Reproducing employment profiles seen in the data

This section evaluates the model's ability to reproduce the life-cycle employment profiles involving self-employment shown in Section 3. Simulating data from the estimated model, I calculate the pairwise distance between the simulated profiles and cluster the simulated profiles. Producing similar profiles and profile-clusters is an informal test of the model's ability to reproduce the rich patterns of self-employment behavior found in the data – moments which are *not targeted* directly in estimation.

Similar to Figure 3, Figure 8 shows the life-cycle employment profiles involving self-employment grouped into seven clusters. The first six clusters are very similar to the clusters found in the data. Specifically, I find clusters consisting of (1) those who enter unincorporated self-employment late in their careers, (2) those who enter incorporated self-employment late in their careers, (3) those who spend most of their careers in unincorporated self-employment, (4) those who spend most of their careers in incorporated self-employment, (5) those who are mostly employed with brief self-employment spells, and (6) those who have weak labor force attachment. The six matched clusters not only reproduce the behaviors seen in the data but also find similar proportions.[36] The seventh cluster consists of individuals who have unincorporated self-employment spells mid-career. This last cluster differs from the seventh cluster found in the data, but in a way that would be expected given that the model groups individuals with similar potential experience while in the data individuals are grouped by age. Specifically, the model involves educational decisions at time $t = 0$ and then simulates years of

[36] Figure 8 also highlights that individuals change careers somewhat less often than in the real data. More accurate rates of career persistence are recovered in a model that assumes non-pecuniary benefits of a particular career depend not only on switching costs, but also acquired experience in that career. For parsimony, I keep the model of career choice which only depends on expected earnings, previous career, the continuation value, and background characteristics.

potential experience, and thus we could not have recovered a group who spent their late 20s in education in the simulated data.

Section E in the Appendix provides diff-in-diff estimates of a policy reform designed to promote small businesses in northern Sweden. I estimate the Diff-in-Diff on the data and on simulated data from the model and show that they produce similar results. This provides additional evidence on the model's ability to fit moments not directly targeted in estimation.

Section G in the Appendix provides further evidence of the model's ability to fit career choices and conditional earnings observed in the data. It shows that the model reproduces the distributions of careers by potential experience and that estimated and observed earnings are very similar.

6.2 Policy counterfactuals

This section of the paper uses the structural model to simulate the various impacts of specific policy changes. Since the model is non-linear and multidimensional, the results section relies on simulations. Specifically, I randomly draw a set of initial regressors from the sample and a latent class from the estimated population distribution, then forward simulate their schooling choices, potential earnings, and career choices.

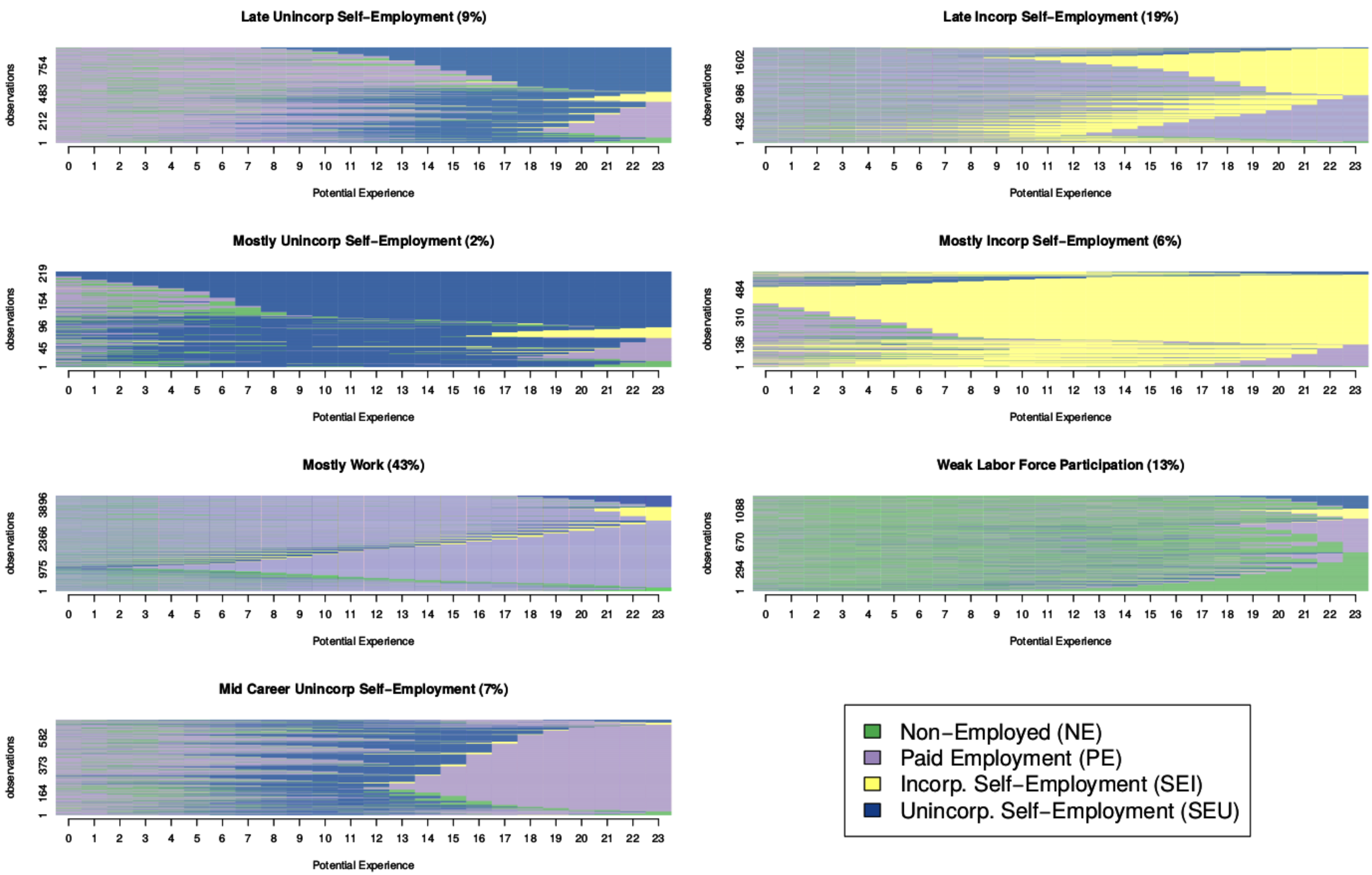
In this section, I consider two sets of policy counterfactuals. The first considers policies designed to change the base-line cognitive and non-cognitive skills of workers. The second considers the surprise introduction of one-time government transfers designed to incentivize individuals to enter self-employment. For the second, I consider both variation in the timing of the policy and in the targeted population.

Changing skills

Many programs aim to equip students with the skills and abilities necessary to be a successful small business owner (see, for example, Elert et al. (2015)). In this section I consider the impact of changing cognitive and non-cognitive skills on self-employment decisions over the life-cycle.[37] Cognitive and non-cognitive skills affect the overall level of self-employment, but

[37] See Heckman and Kautz (2014) for an overview of policies that change cognitive and non-cognitive skills.

Figure 8: Clusters of Simulated Life Cycles Involving Self-Employment.



Notes: Figure shows clusters from running the clustering algorithm described in Section 3 on data simulated from the model.

they also affect the timing and types of self-employment behavior. Policies designed to improve skills will affect educational decisions, earnings, and career decisions in every period.

I evaluate the impact of raising cognitive or non-cognitive skill by one standard deviation. To begin, I summarize the changes in self-employment by estimating how skill changes the self-employment patterns over the life-cycle. To do this, I use the seven clusters from the simulated data shown in Figure 8 and, for each new life-cycle profile, I estimate which of the seven clusters it is most likely belongs to. I assign cluster membership based on the average squared distance between each new life-cycle employment profile and the original seven clusters. This allows me to estimate how much overall self-employment changes, but also how the composition of self-employment behaviors changes.

Table 8 shows how the changes in skill affect the overall rate and composition of self-employment over the life cycle. The first row reports the baseline percent of individuals who are ever self-employed and then the percentage point change in this number from raising cognitive skill and raising non-cognitive skill. The remaining rows show similar results for each of the seven clusters found in the simulated data. Increasing cognitive skill results in an increase in the "Mostly Incorp Self-employment" and "Mostly Unincorp Self-employment" group and decreases the "Weak Labor Force Participation" group. In comparison, raising non-cognitive skills results in a notable rise in the "Mostly Incorp Self-Employment" and "Late Incorp Self-Employment" groups and lowers those in the "Weak Labor Force Participation" and "Mostly Paid Employment" groups.

Another way of evaluating the impact of skill changes is to look at how the various career choices change by age. Table 9 shows how career choices change as a result of the skill increases. The top panel shows results for increased cognitive skill and the bottom panel shows results for increased non-cognitive skill. The first seven columns show the percent change in the various careers and the last column shows the percent change in self-employment overall. The increase in non-cognitive skills substantially increases overall self-employment and self-employment at each age. Interestingly, it increases white-collar jobs in paid employment, incorporated self-employment, and unincorporated self-employment. Increasing non-cognitive skills causes a smaller overall change in self-employment, but also produces different types of self-employment. Increasing non-cognitive skills reduces non-employment substantially more than increasing cognitive skills and raises both blue-collar and white-collar self-employment, except for unincorporated blue-collar self-employment, which remains mostly unchanged. The overall rate of ever being self-employed increases, but not substantially as the reductions in the "Weak Labor Force Participation" and "Mostly Paid Employment" groups reduce the number of individuals who only self-employ one or two years.

Table 8: The Impact of Changing Baseline Skills on Self-Employment Behaviors

		Percentage Point Change	
	Baseline Percent	Cognitive (+1 sd)	Non-Cognitive (+1 sd)
Ever Self-Employed	31%	0.89	0.09
Late Incorp Self-Employment	19.1%	-0.26	1.57
Late Unincorp Self-Employment	9.0%	-0.27	-0.59
Mostly Incorp Self-Employment	6.0%	1.08	2.99
Mostly Unincorp Self-Employment	2.2%	0.82	-0.14
Mostly Paid Employment	43.2%	-0.46	-1.36
Weak Labor Force Participation	13.0%	-1.50	-2.58
Mid-Career Unincorp. Self-Emp	7.2%	0.59	0.10

Notes: This table shows how counterfactual policies raising either cognitive or non-cognitive skill by one standard deviation affect self-employment behaviors. It considers overall levels, as well as the composition of self-employment behaviors. The change in composition is measured by re-classifying each news spell into one of the seven original clusters shown in Figure 8.

Self-employment subsidies by age

This sub-section considers a policy which provides a surprise one-time subsidy for individuals to enter self-employment. I consider this subsidy at ages 24, 29, and 34 to better understand how the timing of the policy affects who is induced into self-employment and how it affects their earnings and utility.

The policy will provide a one-time $10,000 subsidy to those who are self-employed, allowing agents to select into the various forms of self-employment. Table 10 shows how career choices change during first ten years after the intervention. The top panel shows the intervention at age 29, while the bottom panel shows the intervention at age 24. The first seven columns show the percent change in the seven possible career choices while the 8th column shows the percent change in all self-employment.

Table 9: Changes in Career Choices over the Life Cycle from Raising Cognitive and Non-cognitive Skills

	Increased Cognitive Skill							
	Non-Empl	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	Any SE
Perc. Change (total)	-8.0%	7.2%	-9.7%	59.1%	-9.4%	58.1%	-2.8%	24.7%
Perc. Change (age 25)	-9.9%	8.5%	-7.7%	60.1%	-7.4%	68.1%	14.7%	39.1%
Perc. Change (age 30)	-12.0%	7.8%	-9.3%	56.8%	-6.9%	66.5%	4.7%	27.7%
Perc. Change (age 35)	-12.1%	6.9%	-9.4%	48.3%	-8.4%	62.7%	-10.7%	19.9%
Perc. Change (age 40)	-11.9%	7.5%	-10.1%	62.2%	-12.2%	40.3%	-12.3%	16.8%
	Increased Non-Cognitive Skill							
	Non-Empl	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	Any SE
Perc. Change (total)	-18.3%	3.3%	2.5%	26.6%	19.9%	7.6%	-1.4%	12.1%
Perc. Change (age 25)	-17.6%	6.3%	4.8%	21.6%	17.2%	11.2%	-0.9%	9.7%
Perc. Change (age 30)	-18.4%	3.1%	2.4%	22.4%	28.7%	7.6%	3.1%	13.0%
Perc. Change (age 35)	-21.3%	2.4%	1.4%	27.9%	19.3%	5.6%	-0.5%	12.6%
Perc. Change (age 40)	-23.0%	1.7%	2.2%	24.4%	17.9%	3.6%	-3.8%	11.1%

Notes: Table shows how the distribution of careers changes from raising cognitive or non-cognitive skills by one standard deviation. Numbers report what percent that group increased or decreased by as a result of the increased skills.

For the intervention at age 29, in the first year of the intervention, the number of people in self-employment increases by 11%, though this rapidly declines following the intervention, falling to three percent three years after the intervention and 1.6% nine years after the intervention. The various self-employment categories rise by eight to 13 percent in the first year, with the largest increase coming in unincorporated blue-collar self-employment, though this group also has the most rapid decline. The policy initially reduces non-employment by 1.8%, though this fades to a 0.3% reduction in self-employment in the long run. The bottom panel of Table C shows the impact of the same intervention, but implemented at age 24. At age 24, the subsidy causes a much larger percent increase in self-employment which again fades, but not to as low of a level as in the later intervention. Specifically, overall self-employment increases by 30%, fades to 6.4% three years after the subsidy, and 3.2% nine years after the subsidy. The largest initial increase in self-employment again occurs in unincorporated blue-collar self-employment, though this increase again decays rapidly.

Overall, the policies have limited impacts on the earnings and utility of the individuals induced to switch into self-employment. Table 11 provides the percent change in discounted utility and income for those induced into self-employment by the policy. At age 29, the average increase in the present value of utility in dollars is \$4,980, while the average increase in the present value of income is \$9,722. The average overall gain in utility gains is substantially less than the initial \$10,000 subsidy.

Self-employment subsidies targeting specific sub-populations

It is also possible to consider policies that target specific sub-populations. Here I consider three different surprise one-time self-employment subsidies but specifically target different subpopulations.

The first subsidy targets only white-collar self-employment, the second targets only blue-collar self-employment, and the third targets only for those who were non-employed in the previous period. Table 12 shows how career choices change in response to the one-time subsidy. The white-collar subsidy raises self-employment by 4.2% in the year of the subsidy, but this rapidly decays. Those induced into white-collar self-employment are more likely to leave white-collar paid employment and are split evenly between incorporated and unincorporated white-collar self-employment. In contrast, the subsidy for blue-collar self-employment raises total self-employment by 6% and is concentrated in unincorporated self-employment. Finally, the subsidy targeting those who do not work raises overall self-employment by 4% in the year of the intervention and produces mostly unincorporated businesses.

Table 12 shows the average gains in discounted present value of welfare and wages for those induced to change careers by the three policies. The table shows that wage gains are concentrated in those induced into white-collar self-employment and those entering self-employment from non-employment. While the gains are similar in size, they come from different pathways. The gains for those induced into white-collar self-employment come from increased earnings over their previous careers, while the gains for those induced from non-employment come from being more likely to remain in the labor force in the future. While the wage gains are much larger for these two groups, the welfare gains across all three groups remain relatively small, with those induced into white-collar self-employment having the smallest estimated expected gains in utility.

6.3 Careers and self-employment decisions

This section documents the relationship between career, experience, and the costs associated with changing careers. To do this, I consider the returns to different types of experience across careers and the estimated costs of switching careers. Table 14 shows the average returns to an additional year by career across the full population. In the table I consider three types of experience: (1) experience outside the agent's current career and sector, (2) experience in the agent's current sector but not in the current career, and (3) experience in the agent's current career. For each of the three cases, the overall gain is shown with the marginal gains shown in parenthesis.[38]

The table provides three new findings. First, general experience is valued in all careers at approximately the same. Second, experience is more transferable within the white-collar sector than the blue-collar sector. A year of sector-specific (but not career-specific) experience is worth 200 to 400 more dollars than a year of general experience for white-collar workers. In contrast, a year of sector-specific (but not career-specific) experience in blue-collar employment is worth less than a year of general experience, particularly in paid employment. Third, the gains from experience in unincorporated self-employment are low, even within unincorporated self-employment. For both blue- and white-collar unincorporated self-employed workers, the returns to career-specific experience are worth $100 less than sector-specific experience.

Table 15 documents the role of non-pecuniary benefits on career choice and career benefits. The table reports the average total flow utility associated with each career choice. The first row labeled "All" shows this for the full population across the true distribution of previous jobs. The proceeding rows show the distribution of flow utilities when the previous occupation is fixed at a particular previous career.[39] The table provides several insights. First, the estimated utility costs of switching careers are large, particularly for switching into careers involving self-employment. For example, the expected flow utility to entering white-collar incorporated self-employment from white-collar paid employment is estimated to be -$8,621 at age 35, but the flow utility for those already in incorporated self-employment is large and positive. Second,

[38]For case (1), the overall and marginal gains are the same. For case (2), the marginal gains is the additional benefit coming from the experience being in the sector rather than experience outside of the sector. In case (3), the marginal gains are the additional benefits of the experience being in the same career rather than just the same sector.

[39]Note that only previous career is fixed to a counterfactual value and expected earnings in each career are not changed.

the costs of switching careers are smaller for careers in the same sector. For example, the flow utility for white-collar incorporated self-employment is much larger for those coming from white-collar unincorporated self-employment and white-collar paid employment. Similar patterns hold within the blue-collar sector.

6.4 The determinants of capital decisions

This section uses the structural model to quantify how capital decisions are influenced by characteristics of the individual. In particular, I evaluate both how the probability of exiting self-employment the following period and parental wealth affect how much capital individuals employ while self-employed. The first allows me to quantify the importance of future career decisions on capital decisions, while the second provides potential evidence of credit constraints.

Table 16 displays results from a counterfactual simulation where the estimated probability of exit is fixed at 1% and 99%, leaving all other variables unchanged. The table shows the average percent increase in capital from raising the probability of remaining in self-employment from 1% to 99%. The table shows this for each form of self-employment at ages 25 and 35. Across the various forms of self-employment, the amount of capital employed increases by seven to 21% when changing the counter-factual probability of staying in self-employment from 1% to 99%. These differences are smaller than those shown in Table 4 but the simulation controls for the difference in experience, background, and latent type.

Table 17 shows the percent change in self-employment and capital investment from counter-factually fixing parental wealth at the first and 99th percentiles in the population distribution, leaving all other characteristics unchanged. The first column shows the percent change in capital investment among those who self-employ, while the second column shows the percent change in the proportion of the population in that particular career. This counterfactual shift in wealth raises self-employment and investment for all forms of self-employment, but the increase is larger at age 25 than 35 and is larger in the blue-collar sector than the white-collar sector. Parental wealth directly affects the intensive and extensive margin of self-employment decisions. Within the context of the model, this can be interpreted as evidence of credit constraints, but as highlighted by Hurst and Lusardi (2004), it may also be that parental wealth affects preferences for self-employment.

6.5 The importance of persistent unobserved heterogeneity η

This subsection documents that latent heterogeneity plays an important role in the model, even after controlling for the rich set of observable characteristics in the data. Figure 9 displays the distribution of log after-tax earnings for the four different latent classes at one, ten, and 20 years of potential experience. There are small initial differences in the distribution of earnings, most notably the right-shifted earnings of type-3. Twenty years later, there are much larger differences in the earnings distributions between latent classes. Most notably, type-3 individuals have substantially higher earnings while type-2 individuals earn substantially less.

In the estimated model, almost all individuals have posterior probabilities of belonging to a particular type of 0.95 or higher. Assigning individuals to the type with the highest posterior probability, we can characterize career differences by latent type. Most individuals are type-0 or type-1, which account for 40% and 32% of the population respectively. Fifteen percent of the population are the low-earning type-2s, and 12% of the population are the high-earning type-3s. The type-2 and type-3 individuals are also the most likely to be self-employed, with 32% of type-2 and 31% of type-3 individuals being self-employed at some point during their careers. In contrast, 18% and 20% type-0 and type-1 individuals are ever self-employed respectively. Type-2 individuals are the most likely to be non-employed later in their careers and are the most likely to enter unincorporated self-employment, while type-3 individuals are the most likely to ever be in incorporated self-employment. Type-3 individuals do not have substantially higher non-pecuniary benefits from self-employment than the rest of the population, but have notably higher relative earnings in incorporated self-employment. In contrast, type-2 individuals do not have higher relative earnings in unincorporated self-employment, but have larger non-pecuniary benefits associated with unincorporated self-employment.

7 Conclusion

Using population level data from Sweden, this paper provides new evidence on the causes and consequences of self-employment over the life cycle, and evaluates how self-employment decisions can be influenced by policy. To begin, I develop new evidence on the various patterns of self-employment behavior seen over the life cycle, and present an intuitive two-period model to interpret the evidence. Specifically, I use machine learning methods to develop a data-driven taxonomy of self-employment behaviors. Using the taxonomy, I document that there are substantial differences in who selects into different self-employment behaviors.

Motivated and guided by the empirical evidence, I develop a dynamic Roy model where self-employment decisions depend on life-cycle factors such as cognitive and non-cognitive skills, prior work experience, the cost of capital, and other labor market opportunities. The model integrates traditional models of dynamic career choice that feature human capital investment with models of business start-up that feature costly capital investment.

Using the estimated model, I show that the model can reproduce the rich set of self-employment patterns I find in the data. In addition, I find that cognitive and non-cognitive skills, education, and past work experience are important determinants of which types of businesses individuals start, how much capital they employ, and how long they remain in self-employment. Considering counterfactual policies, I find that subsidies designed to promote self-employment are generally ineffective, both in terms of promoting long-lasting firms and in terms of improving the welfare and earnings of those induced to enter self-employment.

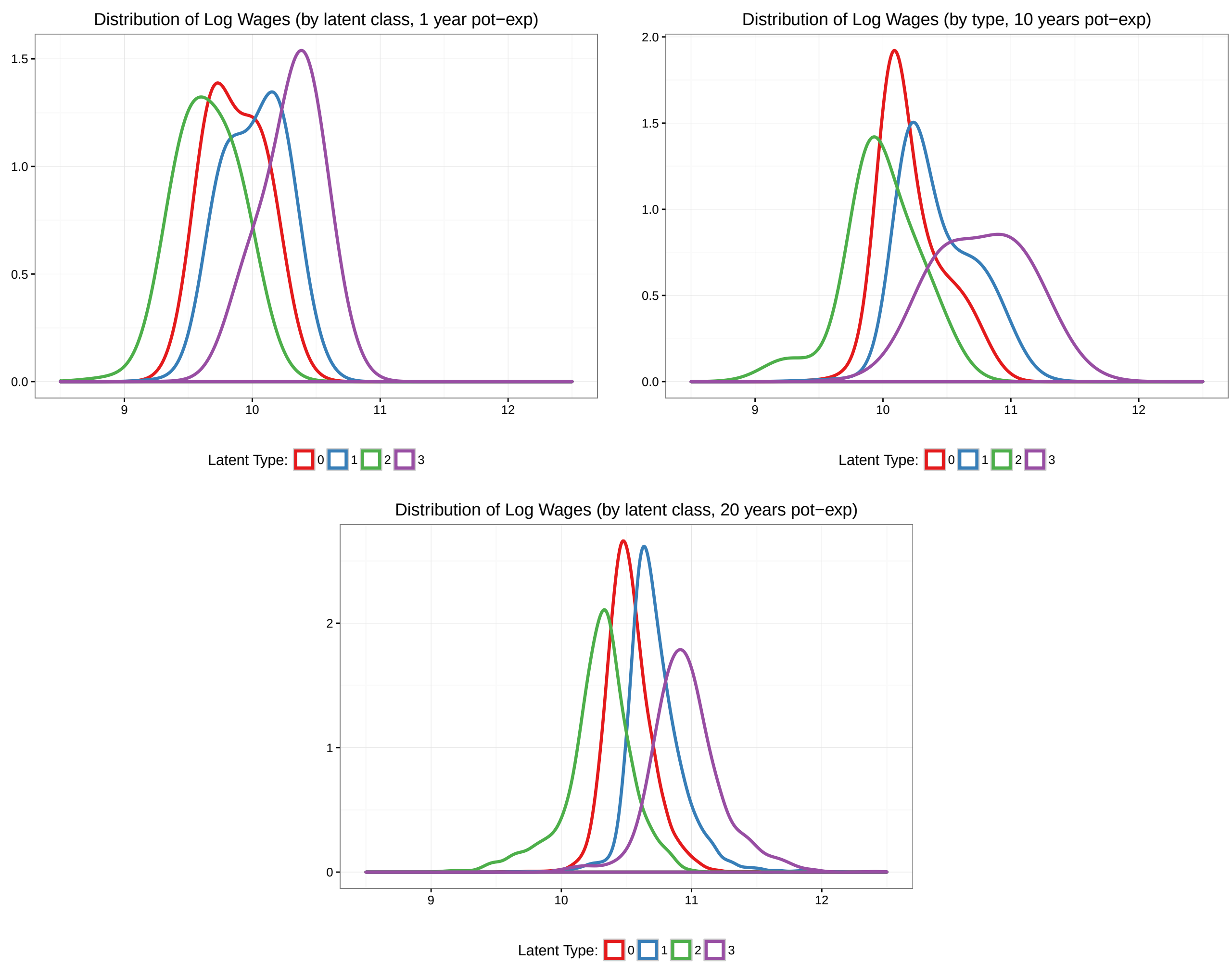
Some caution is in order when interpretting the structural results as I do not model some mechanisms which other work has shown to be important for self-employment and entrepreneurship decisions. First, my model only accounts for risk aversion through differences in non-pecuniary benefits across careers.[40] Second, this paper does not model the private savings of individuals or the capital stock of businesses, both of which may affect self-employment decisions, especially if capital markets are imperfect.[41] Third, the paper does not account for persistent yet time-varying unobservables such as persistent individual-level economic shocks

[40]See Caliendo et al. (2007), Caliendo et al. (2010), Vereshchagina and Hopenhayn (2009), and Herranz et al. (2015) for models and empirical evidence of risk aversion in entrepreneurship and self-employment.

[41]See Evans and Jovanovic (1989), Hurst and Lusardi (2004), Fairlie and Krashinsky (2012),Parker and Praag (2006), Cagetti and De Nardi (2006), and Mondragn-Vlez (2009) for evidence for and against credit constrains and for models that include personal wealth or capital stocks.

or innovations.[42] This paper investigates a number of mechanisms not studied in the literature on self-employment, such as career choice and human capital accumulation, but does not incorporate the mechanisms listed above. In future work, I plan to incorporate more detailed models of the credit market and more flexible risk aversion in order to directly test the empirical relevance of the mechanism proposed here and the more traditional mechanisms considered in the literature.

Figure 9: Earnings Distributions by Experience and Latent Class



Notes: The sub-figures above show the distributions of after-tax income for the four latent classes in the model at one, ten, and twenty years of potential experience.

[42] See, for example, Pakes and Ericson (1998) for a model with technology-specific investment and persistent shocks. See Jovanovic (1982) and Dillon and Stanton (2016) for models where agents learn about entrepreneurial ability over time. Dillon and Stanton (2016) also allow for persistent shocks to wages.

Table 10: Percent Change in Career Choices from Self-Employment Subsidy by Age.

	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	All SE
A: Subsidy for Self-Employment at Age 24								
Age 24	-2.2%	-1.2%	-1.1%	30.4%	29.6%	20.9%	41.6%	30.0%
Age 25	-0.8%	-0.6%	-0.6%	16.2%	17.9%	9.5%	11.70%	12.4%
Age 26	-0.6%	-0.5%	-0.4%	11.1%	11.3%	6.9%	6.8%	8.3%
Age 27	-0.5%	-0.4%	-0.4%	8.8%	9.1%	5.1%	5.20%	6.4%
Age 28	-0.5%	-0.4%	-0.3%	6.7%	7.8%	4.3%	4.70%	5.5%
Age 29	-0.5%	-0.4%	-0.3%	5.4%	7.1%	3.8%	4.00%	4.8%
Age 30	-0.4%	-0.4%	-0.3%	4.4%	5.8%	3.6%	3.80%	4.3%
Age 31	-0.4%	-0.4%	-0.3%	3.6%	4.8%	3.2%	3.60%	3.8%
Age 32	-0.4%	-0.4%	-0.3%	3.3%	4.3%	3.1%	3.20%	3.5%
Age 33	-0.5%	-0.4%	-0.3%	2.8%	3.8%	3.1%	2.90%	3.2%
B: Subsidy for Self-Employment at Age 29								
	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	All SE
Age 29	-1.8%	-0.6%	-0.5%	10.6%	8.0%	9.7%	12.7%	10.6%
Age 30	-0.8%	-0.4%	-0.4%	6.3%	6.4%	5.1%	5.7%	5.8%
Age 31	-0.6%	-0.3%	-0.4%	4.7%	5.0%	3.2%	4.1%	4.2%
Age 32	-0.5%	-0.3%	-0.3%	4.3%	3.8%	2.3%	3.1%	3.3%
Age 33	-0.4%	-0.3%	-0.3%	3.4%	3.1%	1.8%	2.3%	2.6%
Age 34	-0.4%	-0.3%	-0.3%	2.5%	2.5%	1.7%	2.2%	2.2%
Age 35	-0.3%	-0.2%	-0.3%	2.2%	2.2%	1.6%	1.9%	2.0%
Age 36	-0.3%	-0.2%	-0.2%	2.0%	2.0%	1.2%	1.9%	1.8%
Age 37	-0.3%	-0.2%	-0.2%	1.8%	2.0%	1.1%	1.7%	1.7%
Age 38	-0.3%	-0.2%	-0.2%	1.7%	1.7%	1.2%	1.9%	1.6%
C: Subsidy for Self-Employment at Age 34								
	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	All SE
Age 34	-1.3%	-0.5%	-0.4%	4.9%	4.1%	4.1%	5.3%	4.6%
Age 35	-0.6%	-0.3%	-0.3%	3.1%	2.9%	2.4%	2.7%	2.8%
Age 36	-0.4%	-0.2%	-0.2%	2.2%	2.3%	1.7%	1.9%	2.0%
Age 37	-0.5%	-0.2%	-0.2%	1.9%	1.8%	1.4%	1.5%	1.6%
Age 38	-0.4%	-0.1%	-0.2%	1.3%	1.4%	1.0%	1.5%	1.3%
Age 39	-0.4%	-0.1%	-0.1%	1.1%	1.2%	0.9%	1.3%	1.1%
Age 40	-0.4%	-0.1%	-0.1%	0.9%	1.2%	0.8%	1.3%	1.0%
Age 41	-0.3%	-0.1%	-0.1%	0.8%	1.0%	0.7%	1.2%	1.0%
Age 42	-0.3%	-0.1%	-0.1%	0.6%	1.0%	0.9%	1.0%	0.9%
Age 43	-0.1%	-0.1%	-0.1%	0.6%	0.9%	0.7%	1.0%	0.8%

Notes: Table shows the percent change in career choices resulting from a surprise one-time subsidy of $10,000 at ages 24, 29, and 34 to self-employ for the year of the policy intervention and the next nine years. The final column shows the proportion percent increase in all self-employment."SE-I" stands for incorporated self-employment, "SE-U" indicates unincorporated self-employment. "BC" stands for blue-collar and "WC" stands for white-collar.

Table 11: Impact of Subsidies on PV Utility and PV Income

Age of Intervention	Δ PV Utility	Δ PV income	Perc. of Pop Impacted
24	$8,011	$21,268	1.37%
29	$4,980	$9,722	0.76%
34	$729	$10,683	0.50%

Notes: This table shows the wage and utility impacts of a $10,000 subsidy on self-employment at age 24 and 29. Present value of utility and present value of wages are calculated using a discount rate of 0.95. "Perc. of the Pop Impacted" is the percentage of the population induced into self-employment by the policy.

Table 12: Percent Change in Career Choices from Targeted Self-employment Subsidies at Age 29

	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	All SE
A: Subsidy for White-Collar Self-Employment at Age 29								
	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	All SE
Age 29	-0.7%	-0.3%	-0.1%	10.0%	-0.8%	9.5%	-1.0%	4.2%
Age 30	-0.4%	-0.2%	-0.1%	5.9%	-0.2%	5.3%	-0.8%	2.3%
Age 31	-0.3%	-0.1%	-0.1%	3.8%	-0.2%	3.6%	-0.6%	1.5%
Age 32	-0.2%	-0.1%	-0.1%	3.1%	-0.3%	2.5%	-0.5%	1.1%
Age 33	-0.1%	-0.1%	-0.1%	2.4%	-0.2%	2.0%	-0.4%	0.8%
Age 34	-0.1%	-0.1%	-0.1%	2.0%	-0.1%	1.8%	-0.4%	0.7%
Age 35	-0.1%	-0.1%	-0.1%	1.7%	-0.1%	1.7%	-0.4%	0.6%
Age 36	-0.1%	-0.1%	-0.1%	1.6%	-0.1%	1.7%	-0.3%	0.6%
Age 37	-0.1%	-0.1%	-0.1%	1.4%	-0.1%	1.6%	-0.4%	0.6%
Age 38	-0.0%	-0.1%	-0.1%	1.4%	-0.2%	1.5%	-0.3%	0.5%
B: Subsidy for Blue-Collar Self-Employment at Age 29								
	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	All SE
Age 29	-1.1%	-0.2%	-0.4%	-0.9%	8.3%	-0.9%	14.7%	6.0%
Age 30	-0.6%	-0.2%	-0.3%	-0.9%	5.7%	-0.5%	7.5%	3.3%
Age 31	-0.4%	-0.1%	-0.2%	-0.3%	4.3%	-0.4%	5.4%	2.5%
Age 32	-0.3%	-0.1%	-0.2%	-0.3%	3.7%	-0.3%	3.6%	1.8%
Age 33	-0.3%	-0.1%	-0.2%	-0.5%	2.8%	-0.2%	3.1%	1.5%
Age 34	-0.3%	-0.1%	-0.2%	-0.5%	2.6%	-0.1%	2.7%	1.3%
Age 35	-0.2%	-0.1%	-0.2%	-0.4%	2.3%	-0.2%	2.3%	1.1%
Age 36	-0.2%	-0.1%	-0.2%	-0.4%	2.1%	-0.1%	2.2%	1.0%
Age 37	-0.3%	-0.1%	-0.1%	-0.3%	1.8%	-0.2%	2.2%	0.9%
Age 38	-0.3%	-0.1%	-0.1%	-0.3%	1.7%	-0.2%	2.2	0.9%
C: Subsidy for Self-Employment at Age 29 for Those Previously Not Working								
	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC	All SE
Age 29	-1.8%	-0.1%	-0.1%	2.2%	2.9%	4.2%	5.70	4.1%
Age 30	-0.8%	-0.0%	-0.1%	1.6%	1.9%	1.8%	2.10	1.9%
Age 31	-0.6%	-0.0%	-0.1%	1.3%	1.5%	1.0%	1.1%	1.2%
Age 32	-0.5%	-0.1%	-0.1%	1.1%	1.3%	0.6%	0.6%	0.9%
Age 33	-0.4%	-0.0%	-0.0%	0.7%	1.1%	0.7%	0.6%	0.8%
Age 34	-0.4%	-0.0%	-0.1%	0.6%	1.0%	0.5%	0.4%	0.6%
Age 35	-0.3%	-0.1%	-0.0%	0.4%	0.9%	0.5%	0.3%	0.5%
Age 36	-0.3%	-0.0%	-0.0%	0.4%	0.7%	0.3%	0.3%	0.4%
Age 37	-0.3%	-0.0%	-0.0%	0.3%	0.6%	0.3%	0.3%	0.4%
Age 38	-0.2%	-0.0%	-0.1%	0.3%	0.6%	0.4%	0.6%	0.5%

Notes: Table shows the percent change in career choices resulting from a surprise one-time subsidy of $10,000 at age 29. Panel A considers if the policy only targets white-collar self-employment. Panel B considers a policy that targets blue-collar self-employment. Panel C offers the subsidy only to those who were non-employed the previous period. The final column shows the proportion percent increase in all self-employment."SE-I" stands for incorporated self-employment, "SE-U" indicates unincorporated self-employment. "BC" stands for blue-collar and "WC" stands for white-collar.

Table 13: Impact of Subsidies on PV Utility and PV Income

Age of Intervention	Δ PV Utility	Δ PV income	Perc. of Pop Impacted
White-collar SE (age 29)	$4,692	$24,425	0.34%
Blue-collar SE (age 29)	$5,973	$5,053	0.46%
SE from Non-Emp (age 29)	$6,356	$24,892	0.29%

Notes: This table shows the wage and utility impacts of a $10,000 subsidy on self-employment at age 24 and 29. Present value of utility and present value of wages are calculated using a discount rate of 0.95. "Perc. of the Pop Impacted" is the percentage of the population induced into self-employment by the policy.

Table 14: Average Returns to Career-Specific Experience.

	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC
Exp outside sector	$749	$940	$870	$1,005	$680	$970
(marginal gains)	($749)	($940)	$(870)	($1,005)	($680)	($970)
Exp inside sector	$1,129	$655	$1,065	$901	$865	$931
(marginal gains)	($380)	(-$285)	($195)	(-$104)	($185)	(-$39)
Exp inside career	$1,325	$1,051	$2,139	$1,285	$751	$804
(marginal gains)	($195)	($396)	($1,074)	($385)	(-$114)	(-$127)

Notes: This table estimates the average wage gains in the population from a year of additional experience. The first row shows the average gain from an additional year of work experience that is outside of the sector (i.e. white-collar and blue color). The second row shows the average gains from an additional year of experience inside the sector, but not in the specific career. The third row shows the returns to a year of career-specific experience. Below each row the marginal gains are shown in parenthesis.

Table 15: Average Flow Utility by Career (fixing previous career)

Age 25:	Non-emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC
All	15,672	17,117	25,089	-25,888	-2,3298	-28,596	-14,672
Non-employed	**15,672**	4,404	9,255	-35,708	-33,917	-30,612	-19,077
PE White-collar	15,672	**35,341**	15,044	-13,140	-28,909	-11,214	-14,136
PE Blue-collar	15,672	12,475	**38,900**	-32,017	-18,303	-32,415	-10,394
SE-I White-collar	15,672	34,482	19,653	**54,944**	9,128	18,659	-332,259
SE-I Blue-collar	15,672	6,792	24,709	4,464	**46,048**	-339,523	4,426
SE-U White-collar	15,672	15,771	2,383	2,358	-18,072	**31,159**	-4,592
SE-U Blue-collar	15,672	11,030	14,549	-9,062	-5,785	7,440	**31,430**
Age 35:	Non-emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC
All	16,604	29,523	40,351	-17,382	-16,879	-25,851	-11,725
Non-employed	**16,604**	-32,90	-1,112	-37,939	-43,959	-31,343	-25,681
PE White-collar	16,604	**57,455**	22,947	-8,621	-32,996	-4,961	-10,489
PE Blue-collar	16,604	20,234	**64,315**	-25,607	-9,504	-17,319	-129
SE-I White-collar	16,604	43,060	10,434	**66,974**	21,306	6,275	-409,026
SE-I blue-collar	16,604	5,268	35,727	4,087	**62,278**	-41,7381	9,991
SE-U White-collar	16,604	17,703	5,654	6,677	-21,178	**45,131**	-359
SE-I blue-collar	16,604	6,341	19,339	-12,660	7,059	-12,010	**46,718**

Notes: Table shows the estimated average flow utility of each career at age 25 for the full population. The first row shows the estimated flow utilities (in 2010 U.S. dollars) for the full population at age 25. The remaining rows show the flow utilities fixing previous careers for the full population to a particular counter-factual value. This is the marginal effect of changing careers and does not adjust experience or earnings. The utility to non-employment remains constant as earnings are held fixed in this analysis, and period-specific non-pecuniary benefits are normalized to zero in non-employment.

Table 16: Capital and the Probability of Exit in the Future.

Career	Age	%Δ in K
SE-I white-collar	25	14.9%
SE-I blue-collar	25	10.9%
SE-U white-collar	25	12.9%
SE-U blue-collar	25	21.8%
SE-I white-collar	35	7.0%
SE-I blue-collar	35	10.6%
SE-U white-collar	35	11.4%
SE-U blue-collar	35	20.4%

Notes: Table shows the percent increase in capital associated with lowering the probability of exit from self-employment in the next period from 99% to 1%. The averages are taken over the full self-employed population in each age, leaving all other characteristics unchanged.

Table 17: The Relationship between Parental Wealth, Self-Employment, and Capital

Career	Age	%Δ in K	%Δ in Career
SE-I white-collar	25	7.4%	4.7%
SE-I blue-collar	25	16.4%	24.3%
SE-U white-collar	25	10.0%	12.7%
SE-U blue-collar	25	17.7%	19.4%
SE-I white-collar	35	2.5%	0.8%
SE-I blue-collar	35	10.3%	21.5%
SE-U white-collar	35	3.3%	0.5 %
SE-U blue-collar	35	11.2%	10.8%

Notes: The Table shows the percent increase in each form of self-employment and the percent increase in the average amount of capital used in each form of self-employment from raising parental wealth from the 1st percentile to the 99th percentile. The averages are taken over the full population by simulating the full model at both counterfactual values of parental wealth.

References

Abbott, Andrew, "Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, October 1983, *16* (4), 129–147. 00173.

— **and Angela Tsay**, "Sequence Analysis and Optimal Matching Methods in Sociology Review and Prospect," *Sociological Methods & Research*, August 2000, *29* (1), 3–33. 00537.

— **and John Forrest**, "Optimal Matching Methods for Historical Sequences," *Journal of Interdisciplinary History*, 1986, *16* (3), 471. 00329.

Aguirregabiria, Victor and Pedro Mira, "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, May 2010, *156* (1), 38–67. 00264.

Arcidiacono, Peter, "Ability sorting and the returns to college major," *Journal of Econometrics*, July 2004, *121* (12), 343–375. 00000.

— **and John Bailey Jones**, "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 2003, *71* (3), 933–946. 00000.

— **and Robert A. Miller**, "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 2011, *79* (6), 1823–1867. 00000.

Bennmarker, Helge, Erik Mellander, and Bjrn ckert, "Do regional payroll tax reductions boost employment?," *Labour Economics*, October 2009, *16* (5), 480–489. 00052.

Benus, Jacob, Sheena McConnell, Jeanne Bellotti, Theodore Shen, Kenneth Fortson, and Daver Kahvecioglu, "Growing America Through Entrepreneurship: Findings from the Evaluation of Project GATE," Technical Report, IMPAQ International, LLC 2008. 00012.

Blanchflower, David G. and Andrew J. Oswald, "What Makes an Entrepreneur?," *Journal of Labor Economics*, January 1998, *16* (1), 26–60. 00000.

Block, Jrn and Philipp Sandner, "Necessity and Opportunity Entrepreneurs and Their Duration in Self-employment: Evidence from German Micro Data," *Journal of Industry, Competition and Trade*, June 2009, *9* (2), 117–137. 00000.

Bruce, Donald and Herbert J. Schuetze, "The labor market consequences of experience in self-employment," *Labour Economics*, October 2004, *11* (5), 575–598. 00066.

_ and _ , "Tax policy and entrepreneurship," *Swedish Economic Policy Review*, 2004, *2* (11), 233–265. 00122.

Buera, Francisco J., "A dynamic model of entrepreneurship with borrowing constraints: theory and evidence," *Annals of Finance*, April 2009, *5* (3-4), 443–464. 00189.

Cagetti, Marco and Mariacristina De Nardi, "Entrepreneurship, frictions, and wealth," *Journal of political Economy*, 2006, *114* (5), 835–870. 00572.

Caliendo, Marco, Frank Fossen, and Alexander Kritikos, "The impact of risk attitudes on entrepreneurial survival," *Journal of Economic Behavior & Organization*, October 2010, *76* (1), 45–63. 00000.

_ , Frank M. Fossen, and Alexander S. Kritikos, "Risk attitudes of nascent entrepreneursnew evidence from an experimentally validated survey," *Small Business Economics*, December 2007, *32* (2), 153–167. 00000.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda, "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *Journal of Economic Perspectives*, August 2014, *28* (3), 3–24.

Dillon, Eleanor W. and Christopher T. Stanton, "Self-Employment Dynamics and the Returns to Entrepreneurship," *Unpublished*, December 2016. 00000.

Elder, Glen H. and Richard C. Rockwell, "The Life-Course and Human Development: An Ecological Perspective," *International Journal of Behavioral Development*, March 1979, *2* (1), 1–21.

Elert, Niklas, Fredrik W. Andersson, and Karl Wennberg, "The impact of entrepreneurship education in high school on long-term entrepreneurial performance," *Journal of Economic Behavior & Organization*, March 2015, *111*, 209–223.

Ericson, Richard and Ariel Pakes, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 1995, *62* (1), 53–82. 00000.

Evans, David S. and Boyan Jovanovic, "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy*, August 1989, *97* (4), 808–827. 02839.

_ and Linda S. Leighton, "Some Empirical Aspects of Entrepreneurship," *The American Economic Review*, June 1989, *79* (3), 519–535. 02539.

Fairlie, Robert W. and Harry A. Krashinsky, "Liquidity Constraints, Household Wealth, and Entrepreneurship Revisited," *Review of Income and Wealth*, June 2012, *58* (2), 279–306. 00000.

Gabadinho, Alexis, Gilbert Ritschard, Matthias Studer, and Nicalas S. Mller, "Mining sequence data in R with the TraMineR package: A users guide," *Geneva: University of Geneva*, 2009. 00144.

_ , _ , Nicolas S Mller, and Matthias Studer, "Analyzing and Visualizing State Sequences in R with TraMineR," *Journal of Statistical Software*, 2011, *40* (4), 137.

Galichon, Alfred, *Optimal Transport Methods in Economics*, Princeton University Press, 2016.

Goetz, Christopher, Henry Hyatt, Erika McEntarfer, and Kristin Sandusky, "The Promise and Potential of Linked Employer-Employee Data for Entrepreneurship Research," Working Paper 21639, National Bureau of Economic Research October 2015. 00001.

Goetz, Stephan J., Mark Partridge, Steven C. Deller, and David A. Fleming, "Evaluating U.S. Rural Entrepreneurship Policy," *Journal of Regional Analysis & Policy*, 2010, *40* (1), 20–33. 00041.

Haltiwanger, John, *Innovation policy and the economy.*, National Bureau of Economic Research , University Presses Marketing, 2012.

__ **, Ron S. Jarmin, and Javier Miranda**, "Who Creates Jobs? Small versus Large versus Young," *Review of Economics and Statistics*, 2013, *95* (2), 347–361.

__ **,** __ **, Robert Kulick, and Javier Miranda**, "High Growth Young Firms: Contribution to Job, Output, and Productivity Growth," in "Measuring Entrepreneurial Businesses: Current Knowledge and Challenges," University of Chicago Press, 2016.

Hamilton, Barton H., Nidhi Pande, and Nicholas Papageorge, "The Right Stuff? Personality and Entrepreneurship," *Unpublished*, 2015. 00000.

Hamilton, BartonH., "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to SelfEmployment," *Journal of Political Economy*, June 2000, *108* (3), 604–631. 01422.

Heckman, James J. and Tim Kautz, "Fostering and Measuring Skills: Interventions That Improve Character and Cognition," in J. J. Heckman, J. E. Humphries, and T. Kautz, eds., *The Myth of Achievement Tests: The GED and the Role of Character in American Life*, Chicago, IL: University of Chicago Press, 2014, pp. 341–430.

Heckman, James, John Eric Humphries, and Gregory Veramendi, "The Causal Effects of Education on Earnings, Health, and Smoking," *Unpublished Working Paper*, 2016. 00000.

Herranz, Neus, Stefan Krasa, and Anne P. Villamil, "Entrepreneurs, Risk Aversion, and Dynamic Firms," *Journal of Political Economy*, October 2015, *123* (5), 1133–1176. 00002.

Hincapi, Andrs, "ENTREPRENEURSHIP OVER THE LIFE CYCLE: WHERE ARE THE YOUNG ENTREPRENEURS?," *International Economic Review*, 2020, *61* (2), 617–681.

Hopenhayn, Hugo A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, September 1992, *60* (5), 1127. 02296.

Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith, "A Simulation Estimator for Dynamic Models of Discrete Choice," *The Review of Economic Studies*, April 1994, *61* (2), 265–289.

Hotz, V. Joseph and Robert A. Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 1993, *60* (3), 497–529. 00757.

Hurst, Erik and Annamaria Lusardi, "Liquidity Constraints, Household Wealth, and Entrepreneurship," *Journal of Political Economy*, April 2004, *112* (2), 319–347. 00000.

_ **and Benjamin Pugsley**, "What Do Small Businesses Do?," *Brookings Papers on Economic Activity: Fall 2011*, 2011. 00000.

Hurst, Erik G. and Benjamin W. Pugsley, "Wealth, Tastes, and Entrepreneurial Choice," in "Measuring Entrepreneurial Businesses: Current Knowledge and Challenges" NBER Chapters, National Bureau of Economic Research, Inc, 2016.

Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, 1982, *50* (3), 649–670. 04808.

Jr, Glen H. Elder, Monica Kirkpatrick Johnson, and Robert Crosnoe, "The Emergence and Development of Life Course Theory," in Jeylan T. Mortimer and Michael J. Shanahan, eds., *Handbook of the Life Course*, Handbooks of Sociology and Social Research, Springer US, 2003, pp. 3–19.

Jr., Robert E. Lucas, "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, October 1978, *9* (2), 508–523. 02909.

Kasahara, Hiroyuki and Katsumi Shimotsu, "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, January 2009, *77* (1), 135–175.

Keane, MichaelP. and KennethI. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, June 1997, *105* (3), 473–522. 00000.

Levine, Ross and Yona Rubinstein, "Smart and illicit: Who becomes an entrepreneur and does it pay?," *Available at SSRN 2314667*, 2013.

Lindquist, Matthew J., Joeri Sol, and Mirjam Van Praag, "Why Do Entrepreneurial Parents Have Entrepreneurial Children?," *Journal of Labor Economics*, April 2015, *33* (2), 269–296. 00038.

Lindqvist, Erik and Roine Vestman, "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics*, 2011, *3* (1), 101–28. 00000.

Manso, Gustavo, "Experimentation and the Returns to Entrepreneurship," *Review of Financial Studies*, 2016, p. hhw019.

Mondragn-Vlez, Camilo, "The probability of transition to entrepreneurship revisited: wealth, education and age," *Annals of Finance*, January 2009, *5* (3-4), 421–441. 00020.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "The Allocation of Talent: Implications for Growth," *The Quarterly Journal of Economics*, May 1991, *106* (2), 503–530. 00000.

Pakes, Ariel and Richard Ericson, "Empirical Implications of Alternative Models of Firm Dynamics," *Journal of Economic Theory*, March 1998, *79* (1), 1–45. 00000.

Parker, Simon C., *The economics of entrepreneurship*, Cambridge, UK ; New York: Cambridge University Press, 2009. 00000.

_ **and C. Mirjam van Praag**, "Schooling, Capital Constraints, and Entrepreneurial Performance," *Journal of Business & Economic Statistics*, October 2006, *24* (4), 416–431. 00000.

Rust, John, "Chapter 51 Structural estimation of markov decision processes," in "Handbook of Econometrics," Vol. 4, Elsevier, 1994, pp. 3081–3143. 00000.

Schoar, Antoinette, "The Divide between Subsistence and Transformational Entrepreneurship," *NBER*, February 2010, pp. 57–81. 00000.

Simoes, Nadia, Nuno Crespo, and Sandrina B. Moreira, "Individual Determinants of Self-Employment Entry: What Do We Really Know?," *Journal of Economic Surveys*, May 2015. 00000.

Stange, Kevin M, "An Empirical Investigation of the Option Value of College Enrollment," *American Economic Journal: Applied Economics*, January 2012, *4* (1), 49–84. 00000.

Studer, Matthias and Gilbert Ritschard, "What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, February 2016, *179* (2), 481–511. 00005.

Sverige and Riksrevisionen, *Snkta socialavgifter: fr vem och till vilket pris?*, Stockholm: Riksrevisionen, 2008. 00000.

Vereshchagina, Galina and Hugo A. Hopenhayn, "Risk Taking by Entrepreneurs," *American Economic Review*, 2009, *99* (5), 1808–30. 00000.

Vikstrom, Peter, "Entrepreneurship and SME policies across Europe," Working Paper 21, Swedish Agency for Growth Policy Analysis 2011. 00000.

Villani, Cdric, *Topics in Optimal Transportation*, American Mathematical Soc., 2003. Google-Books-ID: idyFAwAAQBAJ.

Vladasel, Theodor, Matthew J. Lindquist, Joeri Sol, and Mirjam van Praag, "On the origins of entrepreneurship: Evidence from sibling correlations," *Journal of Business Venturing*, 2021, *36* (5), 106017.

Wennberg, Karl, Johan Wiklund, Dawn R. DeTienne, and Melissa S. Cardon, "Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers," *Journal of Business Venturing*, July 2010, *25* (4), 361–375. 00000.

APPENDIX

A Policies Aimed at Promoting Entrepreneurship

This section provides details on changes to the tax code that affected self-employment decisions in parts of Sweden. Using this potentially exogenous source of variation I calculate Diff-in-Diff estimates to determine how self-employment decisions, earnings, and hiring choices are affected by the changes in tax rates. I then use the structural model to repeat a more robust version of this exercise.

Since 1990 Sweden has made a number of changes to Sweden's payroll tax. In 1997, social security contribution were reduced by 5 percentage points. It applied to the first 600,000 SEK paid by employers and gradually increased to just under 750,000 SEK. Since the 1990s the tax-rate has remained relatively stable at slighly above 30% (Bennmarker et al., 2009; Sverige and Riksrevisionen, 2008). In 2002, an additional tax break was provided to small businesses operating in Regional Support Area A (RSA A) in Sweden which consists of the roughly the top two thirds of Sweden except for the municipalities on the coast. Figure 10 shows the municipalities included in RSA A. Starting in 1991, this region was granted a 10 percentage point payroll tax reduction, which was reduced to 8 percentage points in 1998 and abandoned in 1999 Bennmarker et al. (2009). Starting in 2002, a new payroll tax was introduced for RSA A which further reduced payroll tax contributions by an additional 10 percentage points for the first 852,000 SEK (approximately $132,000 2010 US dollars) of paid wages.In order to comply with EU regulations, the reduction did not apply to businesses in agriculture or transportation. In 2002, it was announced that the law would stay in place through at least 2007 to minimize uncertainty to potential employers, though, the tax cut was phased out by the end of 2007. In addition, the general 5 percentage point tax cut was reduced to 2.5 percentage points for 2007 and then fully phased out.[43]

Bennmarker et al. (2009) provides a detailed overview of this policy as well as a difference-in-differences analysis of the 2002 change on the hiring decisions of firms, but the analysis focuses on businesses as the unit of analysis and only studies the years 2002 to 2004. In the

[43] In July 2007 a new reduction to payroll taxes was introduced that lowered the payroll tax for employees between the ages of 18 and 25, which in 2009 was expanded to all employees under the age of 26.

simplest specification, they find that a one percentage point reduction in payroll tax raises the gross wage bill by 0.18%, though the relationship is not statistically significant. They similarly do not find a strong effect on the amount of labor used by pre-existing firms, but they do find evidence that the policy induced new firms to enter and that these firms use somewhat more labor than similar pre-existing firms.

Here, I analyse the same set of policies using a Diff-in-Diff strategy to look at how the tax subsidies affected the rate of self-employment, the earnings of the self-employed, and the number of workers employed by the self-employed. For the rate of self-employment and the earnings of the self-employed, I compare diff-in-diff estimates from the data and the simulated data from the structural model laid out in Section 4. RSA A had a ten percentage point reduction in their payroll tax rate from 1991 to 1997 and an eight percent reduction in 1998, which was repealed in 1999, but this policy was replaced by another 10 percentage point reduction for RSA A from 2002 to 2007. Given the policy was in place for most of the 199 to 1999 and from 2002 to 2007, I create two indicator variables for 2000-2001 and 2008-2013 period where 10 percentage point reduction in payroll taxes were not in place for RSA A and interact these indicators with being in the RSA A regions. Using this analysis, I estimate the model:

$$
\begin{aligned}
Y_{i,t} = \alpha_t &+ \gamma_1 \mathrm{RSA}_{i,t} + \gamma_2 \mathrm{RSA}_{i,2000-2001} + \gamma_3 \mathrm{RSA}_{i,2008-2013} \\
&+ \delta_1 \mathrm{Exp}_{i,t} + \delta_2 \mathrm{Exp}^2_{i,t} + \delta_3 S_i + \boldsymbol{X}_{i,t}\beta + \epsilon_{i,t}
\end{aligned}
$$

where $Y_{i,t}$ can be an outcome such as an indicator of self-employment, earnings amoung the self-employed, or number of employees amoung the self-employed. $\alpha_{i,t}$ are year dummies, RSA_i is an indicator of if the individual i lives in RSA A at time t, $\mathrm{RSA}_{i,2000-2001}$ is the RSA indicator interacted with the years 2000 and 2001 which correspond to the first period of no payroll tax reduction for RSA 1, RSA˙i,2008-2013 is the RSA indicator interacted with the years 2008 through 20013 which corresponds to the second period of no payroll tax reduction for RSA 1. $\mathrm{Exp}_{i,t}$ and $\mathrm{Exp}^2_{i,t}$ are potential experience and potential experience squared. S_i is years of schooling and $\boldsymbol{X}_{i,t}$ contains cognitive ability, non-cognitive ability, average parental income, and if the individual's parent's were ever self-employed. In the simulated data, I also control for latent type. When looking at the number of employees, I add industry indicators interacted with t to account for differences in labor intensity across industries and potential industry-specific fluctuations.

In this set-up, γ_2 and γ_3 provide information on the various outcomes were changed in RSA A during the two periods where there was no payroll tax subsidy relative to changes in that time period outside of RSA A. Tables 18, 19, and 20 show the results from the Diff-in-Diff analysis run on all Swedish-born men born between 1965 and 1977. Table 18 shows the baseline impact of living in RSA A on the probability of self-employment, and how that inmpact changes when the payroll tax subsidy is removed. The top panel shows results from the data and the bottom panel shows results from the simulation. The first column shows results for all self-employed workers. On average, in the data, RSA A has one percentage point less self-employment, which is reduced by an additional half of a percentage point during the 2000-2001 reduction in subsidies and by 0.2 percentage points after 2008. For comparison, 8.6% of the total none RSA A population was self-employed in 2004, so the estimated percent reductions from the subsidy are quite large. The second and third columns show the same analysis for incorporated and unincorporated self-employment. During both periods of reduced subsidies, unincorporated self-employment went down by 0.2 percentage points (from a baseline of 4.3% outside of RSA A in 2004). Unincorporated self-employment goes down during the 2000-2001 spell of reduced subsidies, but goes up following the 2008 reduction in subsidies. This may be that those who entered chose to employ fewer workers and lowering the value of incorporation. The results from the simulation are similar to the data except that it estimates a larger increase in unincorporated self-employment in RSA A between 2008 and 2013.

Table 19 shows the impacts of the repealing the subsidies on log income amoung the self-employed. In general, those working in RSA A earn less than those working elsewhere in Sweden, but this is increases when the payroll tax reductions are removed. The top panel shows results on the data while the bottom panel shows results on the simulation. For the subsidy reduction from 2000 to 2002, there is a small but not statistically significant reduction in earnings for all self-employed workers and a 0.02 reduction in log earnings for the incorporated self-employed. For the second reduction of subsidies in 2008 there is a large and persistent reduction in earnings for both incorporated and unincorporated workers. The results from the simulation are similar, though in the simulation the reduction in self-employment income from the removal of the subsidies is larger, both in 2000-2001 and 2008-2013.

Table 18 and Table 19 show that the simulated data produces very similar diff-in-diff estimates as the true data, providing another measure of the structural model's ability to reproduce moments that are not directly targeted in the data.

Table 20 shows the impacts of repealing the subsidies on the number of employees held by businesses. Repealing the subsidies appears to have no impact in 2000-2001, but the 2008 reduction leads to a reduction in employment by 0.2 workers, and a 1 percentage point reduction in the probability of having any additional workers.

Table 18: Diff-in-Diff on Self-Employment from Change in Payroll Tax (comparing estimates from the data and the simulation)

Data:	All SE	Incorp. SE	Unincorp SE
RSA A	-0.009^{***}	-0.002^{***}	-0.008^{***}
	(0.0003)	(0.0003)	(0.0003)
RSA A (2000-2001)	-0.005^{***}	-0.002^{**}	-0.002^{***}
	(0.001)	(0.001)	(0.001)
RSA A (2008-2013)	-0.002^{***}	-0.002^{***}	0.001^{*}
	(0.001)	(0.0005)	(0.0005)
Simulation:	All SE	Incorp. SE	Unincorp SE
RSA A	-0.012^{***}	-0.002^{***}	-0.010^{***}
	(0.001)	(0.0004)	(0.0004)
RSA A (2000-2001)	-0.004^{**}	-0.002	-0.002^{*}
	(0.002)	(0.001)	(0.001)
RSA A (2008-2013)	0.012^{***}	0.001	0.011^{***}
	(0.001)	(0.001)	(0.001)

Notes: Table shows Diff-in-Diff estimates on the proportion of individuals self-employed. The top panel shows results for the data and the bottom panel shows results from data simulated from the structural model laid out in Section 4. The first column pools the self-employed, while the second column is only on the incorporated self-employed and the third column is only on the unincorporated self-employed. Controls for background characteristics, education, and skill are not displayed.

It is also possible to evaluate this policy change within the structural model developed in this paper. Specifically, I allow indicators for living in RSA1 to enter the choice equation each period and allow an indicator for living in RSA A to enter the conditional earnings equations.

Table 19: Diff-in-Diff on Log Earnings from Change in Payroll Tax (comparing estimates from the data and the simulation)

Data:	All SE	Incorp. SE	Unincorp SE
RSA A	-0.033^{***}	-0.079^{***}	-0.043^{***}
	(0.004)	(0.005)	(0.007)
RSA A (2000-2001)	-0.007	-0.026^{**}	0.012
	(0.013)	(0.013)	(0.020)
RSA A (2008-2013)	-0.101^{***}	-0.062^{***}	-0.100^{***}
	(0.007)	(0.007)	(0.011)
Simulation:	All SE	Incorp. SE	Unincorp SE
RSA A	-0.046^{***}	-0.091^{***}	-0.038^{***}
	(0.002)	(0.002)	(0.002)
RSA A (2000-2001)	-0.053^{***}	-0.034^{***}	-0.067^{***}
	(0.006)	(0.005)	(0.005)
RSA A (2008-2013)	-0.145^{***}	-0.067^{***}	-0.189^{***}
	(0.004)	(0.003)	(0.003)

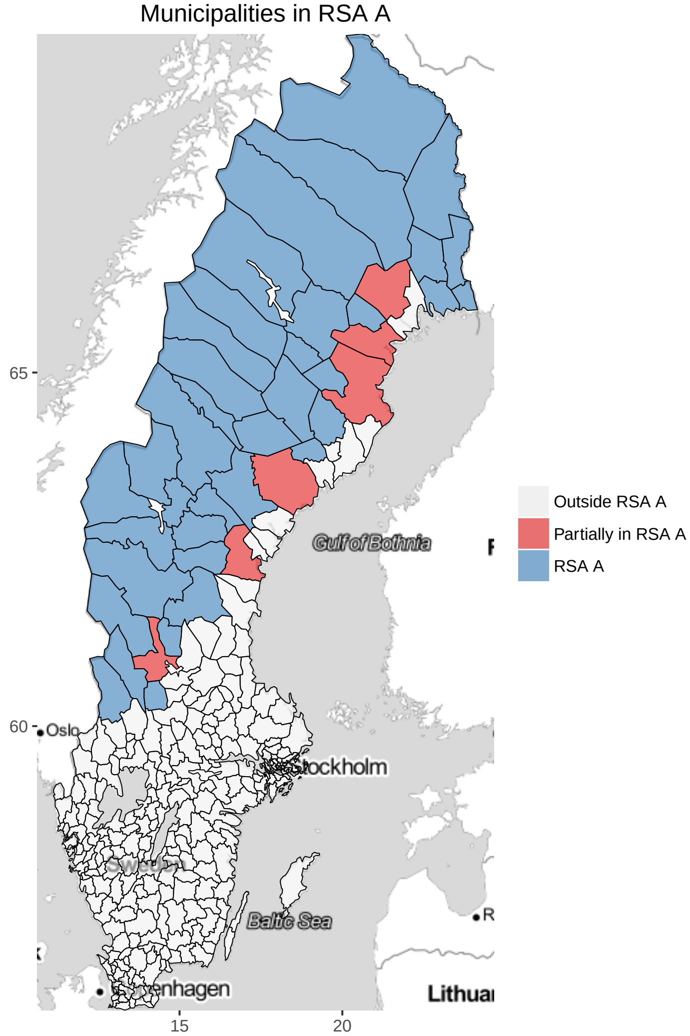
Notes: Table shows Diff-in-Diff estimates on the log after-tax income fo the self-employed. The top panel shows results for the data and the bottom panel shows results from data simulated from the structural model laid out in Section 4. The first column pools the self-employed, while the second column is only on the incorporated self-employed and the third column is only on the unincorporated self-employed. Controls for background characteristics, education, and skill are not displayed. *p<0.1; **p<0.05; ***p<0.01

Table 20: Diff-in-Diff on Number of Employees from Change in Payroll Tax

	Employees	Any Employees	Employees (incorp)	Employees (unincorp)
RSA A	0.084	0.052^{***}	-0.773^{***}	0.490^{*}
	(0.079)	(0.003)	(0.248)	(0.271)
RSA A (2000-2001)	0.075	0.0001	-0.111	1.087
	(0.223)	(0.008)	(0.686)	(0.741)
RSA A (2008-2013)	-0.202^{*}	-0.012^{***}	0.074	-0.296
	(0.117)	(0.004)	(0.357)	(0.409)
Observations	873,553	820,709	420,142	387,663

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 10: Map of Municipalities Affected by Payroll Tax Cut



Notes: Figure shows the municipalities which are part of RSA A which was subject to a 10 percentage point reduction in payroll taxes. The six municipalities shown in red have a portion which is included in RSA A. For the costal municipalities, the coastal area is not included in RSA A.

In this way, I can estimate how living in RSA A affected career choices and outcomes in each period. These coefficients are year-specific, but we can average these coefficients over the periods in which there were no payroll tax subsidies and compare them to the average of the coefficients on RSA1 in the years there were subsidies. Moreover, the structural model accounts for previous choices and latent heterogeneity of the agents, potentially providing a more-robust estimate of the impact of the policy change on self-employment behavior.

B Additional Details on the Data and Institutional Setting

This analysis uses Swedish administrative data for men born between 1965 and 1977 who successfully completed the mandatory military service exam. By linking several different administrative data sets we are able to build detailed background and labor market histories. The data include detailed information on parental background; labor market participation and performance; educational enrollment, attainment and performance; records from mandatory military exams at age 18 or 19 that include measures of cognitive ability, stress tolerance, BMI, and baseline health; and detailed information on the businesses started by people in our sample. This section focuses on documenting the various data sets used in the analysis.

This analysis uses the Longitudinal integration database for health insurance and labour market studies (LISA), a linked register database maintained by Statistics Sweden which extends previous versions of the registry data used for labor market research (LINDA and LOUISE).[44] The data combines a number of Swedish registers starting in 1990. Specifically they collect data on "labor market, educational and social sectors." The data focuses on the individuals, but provides links between family members and employers. This research uses three different portions of the LISA data: (1) individual records which provide detailed income accounts, employment, and education. (2) Company-level data on number of employees, current assets, and fixed assets. (3) Establishment level data on industry, number of employees, and employee compensation.

The data merges together detailed individual tax and administrative records from the LISA data sets from Statistics Sweden that includes detailed earnings data from wages, asset income, active and passively owned businesses income, income from the government for education, disability income, income for taking care of young children or elderly family members, and pension income. LISA data includes basic educational attainment data, the sector and occupation of employment, and basic details about their employer, such as if the individual is employed by the government, or by a privately owned company. These data also allow us to know when individuals change employers, worksites, occupations, and industries.

[44] See http://www.scb.se/lisa-en for additional details.

The LISA data also includes detailed information on immigration status, including country of origin. The LISA data is available for the individuals in our sample, as well as their siblings and parents.

Using the baseline LISA data, a number of additional administrative data-sets are linked:

- **9th Grade Education Data:** Detailed records from 9th grade allow us to see where the individual attended the last year of mandatory schooling, their 9th grade GPA.
- **High School Education Data:** Details on if the individual enrolled in high school, which track they enrolled in, what their focus was, and their GPA.
- **Post-secondary Education Data:** Enrollment and completion of post-secondary and vocational courses.
- **Military Exam Data:** Every male Swedish citizen in the cohorts had required conscripted service in the military. To aid assigning individuals to tasks, all males are required to attend a day of military testing. For each male in our sample, we have (1) measures of cognitive ability, (2) aggregate measure from evaluation by military psychologist on ability to handle the rigors of active military duty. Based on verbal accounts of what the aggregate score attempts to measure, it would a combination of stress tolerance, conscientiousness, and a propensity for leadership, and (3) BMI and basic health measures. See Lindqvist and Vestman (2011) for an overview on the measures collected during the military examinations.

Self-employment data Using the combined data set, I have a unique picture of the role of self-employment over the life cycle:

- **Human Capital** We have detailed measures of early-career skills and human capital including cognitive ability, stress tolerance, 9th grade GPA, high school track and focus, and college and vocational training – including choice of major. The data also has detailed experience variables, such as years of labor market experience, years of self-employment experience (by industry), and firm specific tenure.The data also contains details about the size of the firm at which individuals are employed, in case firm size affects the type of human capital acquired on the job.

- **Incorporation** We know if self-employed individuals are incorporated or not. If they are incorporated, we know if ownership is shared with other individuals.
- **Capital, and Profits** For each company we have detailed financial records including reported total assets, current assets, fixed assets, number of employees, and employee compensation.
- **Additional firm and work-place information** For each company we know their current assets and liabilities, including long-term liabilities that may affect exit in the short run. We also know the total number of employees, total wage bill, and average educational level of employees. We have data both for the firm and workplace.

  Income data from LISA The data merges together detailed individual tax and administrative records from the LISA data sets from Statistics Sweden that includes detailed earnings data from wages, asset income, active and passively owned businesses income, income from the government for education, disability income, income for taking care of young children or elderly family members, and pension income. Other less detailed records on total income and disposable income al

- **Exposure to Entrepreneurship** We know if parents or siblings are ever self-employed, as well as the prevalence of self-employment in their region of residence.
- **Detailed control variables** In addition to the variables mentioned above, the data allows us to control for parental income when young, parental education, the age of the mother at time of birth, region of birth, number of siblings, birth order, local economic conditions, number of young children in the household, and employment status and occupation of other family members.

Details on Self-employment in Sweden Since reforms in the early 1990s, Sweden's tax system has set up such that there are few incentives for businesses to hire workers as private contractors rather than employees. This is further reinforced as, unlike the Unite States, the social safety-net is not tied to employers. The social safety-net is larger in Sweden than in the United States, but these benefits (and the associated tax wedge) are similar to other Western European countries. In order to qualify for benefits associated with job-loss, the owner must prove that they have fully disbanded their business, preventing small businesses from using unemployment benefits as a way to smooth over periods of slow business.

As in the United States, small businesses can be incorporated or unincorporated. Incorporation requires a one-time registration fee of $270 USD and the business must demonstrate that they own or employ $6,750 USD of capital. As in the United States, there are tax advantages from incorporation for large revenue businesses. Similar to the United States, incorporation better protects personal assets, but not as strongly as in the US. Bankruptcy laws tend to be less lenient on businesses and individuals than in the US, but are more lenient than many other countries in Europe.

Unemployment benefits in Sweden Unemployment benefits in Sweden are based on earnings over the previous three years and the cover up to 80% of prior wages (up to $30,300 USD). In order to qualify for benefits, individuals must have worked at least 80 hours a month for the last six months. The benefits reduce to 70% after 200 days and 65% after 300 days. Continued unemployment benefits may require the unemployed to accept job offers, participate in training programs, or participate in subsidized part-time work. Refusal of job offers or refusal to participate can result in the loss or reduction of benefits. If the self-employed enter unemployment after a short spell, they can opt to use their prior employment for the determination of benefits.

Education in Sweden Post-secondary education is highly subsidized and students tend to spend more time in post-secondary education than in the United States. As with most European Countries, students are tracked and choose a focus in high school. Similarly, students choose a focus of study prior to entering a post-secondary degree. Vocational degrees (secondary and post-secondary) are more common than in the United States, but are less common than in other European countries such as Germany.

C Comparing the Swedish and US Economies

At first glance, an analysis of self-employment in Sweden may not seem to generalize to other settings. This belief is based largely on the notion that Sweden's economy remains similar to the earlier, more highly regulated, economy of the 1970s and 1980s. Following a number of large reforms to the welfare state, tax code, and business regulations in the early 1990s, Sweden has been held up as an example of how economic policies, combined with fiscal reform, can promote innovation and entrepreneurship in Europe. Sweden is currently the top-ranked country on the European Union's innovation scoreboard and ranked similarly to the United States in terms of the ease of doing business.[45] Sweden provides a larger social safety net than the United States through universal health care, more generous unemployment benefits, and higher rates of unionization, but these (as well as the country's tax wedge) are similar to other countries in Western Europe. However, unlike the United States and some European countries, the Swedish tax code provides few incentives for businesses to pay workers as private contractors rather than employees. Given Sweden's success at promoting new growth-oriented businesses, Europe's interest in adopting similar policies, and the similarity between the Swedish and U.S. economies, Sweden serves as an ideal laboratory for evaluating policies designed to promote self-employment and entrepreneurship.

D Details on Trajectory Analysis and Sequence Dissimilarity Measures.

Looking at a person's "state" over time was first considered by Abbott (1983) and has developed into a broader literature on trajectory analysis (Abbott and Forrest, 1986; Abbott and Tsay, 2000). This section provides details on trajectory analysis and how dissimilarity measures are calculated.

Starting with Abbott (1983), trajectory analysis has aimed to better understand the importance of experienced events as well as their duration, order, and timing over the life cycle. raTrajectory analysis constructs sequences of discrete states for individuals. These sequences contain information about which states people experience, the distribution of states, the timing of states, the duration of specific state spells, and the sequencing of states. As pointed out by Studer and Ritschard (2016), these five types of information are not independent and we can

[45] Sweden ranks eighth on the World Bank's ease-of-doing-business rankings (one rank behind the United States), and ranked sixth in The Economist's business environment ranking (one rank ahead of the United States). Moreover, Sweden has observed the second largest number of new companies per capita with valuations over one billion dollars since 2003, behind only the U.S.

typically focus on the sequence, duration, and timing of states over individuals' lives.

In order to compare two states, researchers have developed a number of metrics for constructing dissimilarity metrics between two sequences of discrete states. Studer and Ritschard (2016) provides a review and comparison of 18 different dissimilarity measures. As they highlight, these can be grouped into three categories: distance between distributions, counts of attributes common between two distributions, and edit distance. Choosing between the various forms of distance places more or less emphasis on timing and sequence of events. For example, counting the number of states that are the same within each period makes the measure of distance very sensitive to timing, while being less sensitive to duration or sequence. In comparison, counting the total time periods spent in each state does not depend on timing or duration of any particular spell.

The most common approach in empirical exercises has been "Optimal Matching" which is a measure of edit distance that calculates how many insertions, deletions, and substitutions are needed to transform one string into the other string S_2. Associated with each of the three operations is a cost and the distance between two strings is the cost for the cost-minimizing set of transformations that make the two strings identical. As the costs of insertion and deletion are made relatively lower than the cost of substitution, the model allows for more flexible "time warping" – placing more emphasis on sequencing and less emphasis on timing.

Let $\boldsymbol{D}$ be the matrix of Optimal-Matching distances between every pair of trajectories. Using this dissimilarity matrix, it is then possible to cluster life trajectories using any standard hierarchical clustering algorithm. For example, the commonly used "Ward's Method," starts with each observation being in its own cluster and sequentially looks for the two clusters to merge that minimize the increase in total within-cluster variance. This algorithm produces a tree of potential groupings, and the researcher is left to choose how many clusters to use in practice. Ward's method is one criteria for constructing a tree of groupings from the distance matrix. Other common approaches for determining which clusters should be merged are maximum distance, minimum distance, and average distance between elements within each cluster.

E Assumptions Related to the Dynamic Discrete Choice Model

The dynamic discrete choice model used in this paper relies on a standard, though admittedly strong, set of assumptions common in the dynamic discrete choice literature. These assumptions are laid out in Table 21. These assumptions follow closely to the assumptions made in Rust (1994) and Aguirregabiria and Mira (2010) with the exception that I allow for the unobservable state variable η to be known by the agent and to enter non-linearly into the choice equation and outcome equations. This is similar to the latent types used in Keane and Wolpin (1997) and Arcidiacono and Jones (2003). Similar to Heckman et al. (2016), I assume that the correlation between career decisions and earnings is captured by observable state variables as well as an unobservable latent heterogeneity η, but that the idiosyncratic earnings shocks $\xi(d_t)$ and the idiosyncratic shocks to choices $\epsilon(d_t)$ are independent across choices, individuals, and time periods (assumptions IID and CIX).[46] I assume agents make career decisions based on their expected utility given their observable characteristics and η. This implies that agents do not know $\boldsymbol{\xi}$ when making career choices and the utility function entering the choice equation can be written as $\int U(\boldsymbol{a}_t, \boldsymbol{Z}_t, \eta, \xi(d_t))dG_{\xi(d_t)}(\xi(d_t)) + \epsilon_t(d_t) = u(\boldsymbol{a}_t, \boldsymbol{Z}_t, \eta) + \epsilon(d_t)$, where the second equality defines $u(\ )$ as the expected utility when η and $\epsilon(d_t)$ are known but $\xi(d_t)$ is unknown. Assumptions IID and CIX imply that we can write the Markov state transitions $F_{t+1}(\boldsymbol{Z}_{t+1}, \boldsymbol{\epsilon}_{t+1}, \boldsymbol{\xi}_{t+1}|a_t, Z_t, \eta, \boldsymbol{\epsilon}_t, \boldsymbol{\xi}_t)$ as $F_{Z,t+1}(\boldsymbol{Z}_{t+1}, \eta|\boldsymbol{a}_t, \boldsymbol{Z}_t, \eta)G_\epsilon(\boldsymbol{\epsilon}_{t+1})G_\xi(\boldsymbol{\xi}_{t+1})$.

Using these assumptions, we can rewrite Equation 9 as

$$
\begin{aligned}
V_t(\boldsymbol{Z}_t, \eta, \boldsymbol{\epsilon}_t) = \sum_{\boldsymbol{a}_t \in A} \mathbb{1}\{\delta_t(\boldsymbol{Z}_t, \eta, \boldsymbol{\epsilon}_t)) = \boldsymbol{a}_t\} \Big[& u(\boldsymbol{a}_t, \boldsymbol{Z}_t, \eta) + \epsilon(a_t) \\
& + \beta \int \bar{V}_{t+1}(\boldsymbol{Z}_{t+1}, \eta) dF_t(\boldsymbol{Z}_{t+1}|\boldsymbol{Z}_t, \eta) \Big] \\
= \sum_{\boldsymbol{a}_t \in A} \mathbb{1}\{\delta_t(\boldsymbol{Z}_t, \eta, \boldsymbol{\epsilon}_t) = \boldsymbol{a}_t\} & v_t(\boldsymbol{a}_t, \boldsymbol{Z}_t, \eta, \epsilon(a_t))
\end{aligned}
\tag{15}
$$

where $\bar{V}_t = \int V_t(Z_t, \eta_t) g(\epsilon_t) d\epsilon_t$ is the integrated value function and $v_t(a_t, Z_t, \eta_t, \epsilon_t)$ is the choice-specific value function, where I have assumed optimal K_t can be expressed as a function of Z_t

Table 21: Baseline Modeling Assumptions

PAS	partial additive separability	The flow utility is separable in one of the two unobserved components such that $U_t(\boldsymbol{a}_t, \boldsymbol{\Omega}_t) = u_t(\boldsymbol{a}_t, \boldsymbol{Z}_t, \eta, \boldsymbol{\xi}_t) + \epsilon(\boldsymbol{a}_t)$, where, similar to the assumption of a random utility model, we assume $\mathrm{E}(\epsilon(a_t)) = 0$ and $\epsilon(a_t) \in \mathbb{R} \; \forall \; a_t \in A$.
IID	iid unobservable	$\boldsymbol{\epsilon}_t \in \tilde{\Omega}_t$ is independently and identically distributed across time and individuals and has CDF G_ϵ which is assumed to have finite first moment, to be continuous, and to be twice differentiable. Similarly, $\boldsymbol{\xi}_t$ is assumed to iid within individuals in a time period with CDF $G_{\xi,t}$. ξ_t is assumed be continuously distributed on $\mathbb{R}^+$ with $E[\xi_t] = 1$.
CIX	conditional independence of future $\boldsymbol{Z}_{t+k}$	Conditional on the action $\boldsymbol{a}_t$, current state variable $\boldsymbol{Z}_t$ and current non-additive error η, the cumulative distribution function $F_{\boldsymbol{Z},t+1}(\boldsymbol{Z}_{t+1}, \vert \boldsymbol{a}_t, \boldsymbol{Z}_t, \eta, \boldsymbol{\xi}_t, \boldsymbol{\epsilon_t})$ can be written as $F_{Z,t+1}(Z_{t+1}, \vert a_t, Z_t, \eta)$.
CIO	conditional independence of outcomes	At the time the discrete choice d_t is made, the shock $\xi(d_t)$ associated with earnings outcomes are unknown to individual. When agents make occupational decisions d_t, their information set includes state-variables $\boldsymbol{Z}_t$, $\boldsymbol{\epsilon_t}$, and η.
MLOGIT	mixed logit distribution assumptions	Similar to a mixed logit model, I assume that $\epsilon(d_t)$ are iid with a type-1 extreme value distribution and that η are iid with discrete probability mass function f_η.
DET-K	capital does not depend on idiosyncratic shocks	Agents do not know $\xi(d_t)$ when making investment decisions, so capital can be written as $K_t = K(d_t, \boldsymbol{Z}, \eta)$

Notes: It is possible to relax the final two assumptions, though I do not do so in this paper. Many of these assumptions are standard in the dynamic discrete choice literature, see Rust (1994) and Aguirregabiria and Mira (2010) for an overview. Where possible, I use similar notation and descriptions as in Aguirregabiria and Mira (2010) which surveys the literature.

and η_t and substituted out of the maximization problem.

F Additional Details on Estimation

This section provides additional details about the use of conditional choice probabilities and finite dependence used to recover the estimated choice-specific integrated value function, which serves as an offset parameter in the period-specific multinomial career choice decisions.

[46] In addition it is straight forward to allow the second moment of $\xi(d_t)$ to depend on η.

F.1 Conditional choice probabilities and finite dependence.

Starting with Hotz and Miller (1993), there has been a large amount of work on estimating structural models taking advantage of an alternative representation of the Bellman equation. Specifically, under the MLOGIT assumptions laid out in Table 21, we can express the probability of an particular choice as:

$$p_t(j|\boldsymbol{Z}_t, \eta) = \frac{exp\left(\bar{v}_t(j, \boldsymbol{Z}_t, \eta)\right)}{\sum_{j' \in D} exp\left(\bar{v}_t(j', \boldsymbol{Z}_t, \eta)\right)} \tag{16}$$

where $\bar{v}_t(j, \boldsymbol{Z}_t, \eta)$ is integrated choice-specific value function of a person choosing career j at time t with state variables $\boldsymbol{Z}_t$ and η. Next, using the results from Hotz and Miller (1993), the integrated value function can be written as:

$$\begin{aligned} \bar{V}_t(\boldsymbol{Z}_t, \eta) &= \delta + ln\left(\sum_{j' \in D} exp\left(\bar{v}_t(j', Z_t, \eta)\right)\right) \\ &= \delta + ln\left(exp\left(\bar{v}_t(j^*, Z_t, \eta)\right)\left[\frac{\sum_{j' \in D} exp\left(\bar{v}_t(j, Z_t, \eta)\right)}{exp\left(\bar{v}_t(j^*, Z_t, \eta)\right)}\right]\right) \\ &= \bar{v}_t(j^*, Z_t, \eta) - ln(p(j^*|Z_t, \eta) + \delta \end{aligned} \tag{17}$$

where δ is Euler's constant and $p(j|Z_t, \eta)$ is the probability a person with state variables Z_t and η choose option j^*. One of the key insights of Arcidiacono and Miller (2011), is that we can write the integrated value function as the value of a particular choice j^*, an adjustment term for the fact choice j^* may not be the optimal choice (which is only a function of the conditional choice probability, and Euler's constant). The key insight here is that j^* can be any choice and need not be the optimal choice at time t.

In this representation, the future integrated value function has been replaced with the integrated choice specific value function for a choice of our choosing. Given the model is finite, we can recursively substitute in for one particular path to time T, requiring us to theoretically only evaluate one potential path of choices, greatly reducing the computational time needed to estimate the model. In practice, usually several different paths are simulated and averaged to minimize estimation error(See Hotz et al. (1994)).

Finite Dependence A model exhibits finite dependence if a particular choice at time t and the baseline choice at time t (in this case non-employment) have particular future paths that result in arriving at the same expected state some number of periods in the future. Finite dependence further simplifies estimation, but also imposes strong assumptions on the model. In practice, we estimate utilities in discrete choice problems relative to a reference choice, which here will be non-employment. Arcidiacono and Miller (2011) show that if two different initial choices can lead to the same state in expectation, it is only necessary to perform the recursion above until the that point is reached. In the context of this paper, state variables are assumed to be fixed traits of the individual, economic conditions that are independent of previous choices, experience vector E_t which evolves in a deterministic manner, and previous one period of career and investment choices a_{t-1}. Thus, if a person chose the sequence of choices $\{j_t = se, j_{t+1} = ne, j_{t+2} = pe\}$, and assuming experience does not depreciate, they would reach the same state in expectation as if they had chosen: $\{j_t = ne, j_{t+1} = se, j_{t+2} = pe\}$. Thus, we can express choice-specific value functions as:

$$\begin{aligned} v_t(j, Z_t, \eta, \epsilon_t) =& u_t(j, Z_t, \eta) + \epsilon_t(j) \\ &+\beta \int \Big(u_{t+1}(j^*_{t+1}, Z_{t+1}, \eta) - ln(p(j^*_{t+1}|Z_{t+1}, \eta) \Big) dF_t(Z_{t+1}|Z_t, \eta) + \beta\delta \\ &+\beta^2 \int\int \Big(u_{t+2}(j^*_{t+2}, Z_{t+2}, \eta) - ln(p(j^*_{t+2}|Z_{t+2}, \eta) \Big) dF_{t+1}(Z_{t+2}|Z_{t+1}, \eta) dF_t(Z_{t+1}|Z_t, \eta) + \beta^2\delta \\ &+\beta^3 \int\int\int \Big(\bar{V}_{t+3}(Z_t, \eta) \Big) dF_{t+2}(Z_{t+3}|Z_{t+2}, \eta) dF_{t+1}(Z_{t+2}|Z_{t+1}, \eta) dF_t(Z_{t+1}|Z_t, \eta) + \beta^3\delta \end{aligned} \tag{18}$$

Now, by choosing the correct sequence of future choices,

$$v_t(j, Z_t, \eta) - v_t(0, Z_t, \eta)$$

will only require calculating flow utilities and choice probabilities for two periods into the future as the fourth line of Equation 18 will be identical and cancel.[47] This further simplifies calculations as $\beta \int \bar{V}_{t+1}(Z_{t+1}, \eta) dF_t(Z_{t+1}|Z_t, \eta)$ can be estimated using only conditional choice probabilities, flow utilities, and state transitions two periods into the future.

[47]Note this only holds if previous career trajectory is fully summarized by non-depreciating human capital stocks and an indicator of previous occupation as laid out in Section 4.

The Likelihood Estimation assumes that observable covariates $\boldsymbol{Z_t}$ and persistent unobserved state η generate the dependence across outcomes. In addition, conditional on expected outcomes, covariates $\boldsymbol{Z}_t$ and persistent unobserved state η generate dependence between career decisions and outcomes. If both $\boldsymbol{Z}_t$ and η were observed, the estimation can be decomposed into a number of separable maximum likelihood estimations. Here I will lay out estimation assuming η is known. In the next subsection I lay out how to use the EM algorithm to recover the distribution of η through an iterative process.

The overall likelihood is given by:

$$\mathcal{L} = \prod_i f(\boldsymbol{Y}_i, \boldsymbol{d}_i, \boldsymbol{K}_i, s_i) \tag{19}$$

$$\tag{20}$$

where $\boldsymbol{Y}_i$ is the full set of outcomes, $\boldsymbol{d}_i$ is the full set of career decisions, $\boldsymbol{K}_i$ is the full set of capital decisions, s_i is the choice of education, and $f(\ )$ is the probability density function. This likelihood can be decomposed into a number of small recursive estimations. First, using the set of people self-employed in the final period ($h_T = se$), we estimate :

$$\mathcal{L}_T^K = \int \prod_i \prod_{\delta \in D} \left(f_\delta(K_T|\boldsymbol{Z}_T, s, \eta)\right)^{\mathbb{1}\{d_T=\delta\}} f(\eta)d\eta \tag{21}$$

$$\mathcal{L}_T^Y = \int \prod_i \prod_{\delta \in D} \left(f_\delta(Y_{i,T}|\boldsymbol{Z_T}, \bar{K}_{d_T}, s_i, \eta)\right)^{\mathbb{1}\{d_T=\delta\}} f(\eta)d\eta \tag{22}$$

$$\mathcal{L}_T^d = \int \prod_i f(d_{i,T}|\boldsymbol{Z}_T, \bar{\boldsymbol{Y}}_T, s_i, \eta) f(\eta)d\eta \tag{23}$$

where we assume the optimal choice of capital when not self-employed is zero.

Once period T has been estimated, for period $T-1$ we construct the forward looking elements needed for choice of capital at time $T-1$ and to construct estimates of the various choice-specific continuation values. First, I non-parametrically estimate the probability of remaining self-employed one period in the future for each self-employment state ($P(d_{t+1} = d_t|\boldsymbol{Z}_t, \eta)$, then I estimate $\bar{K}_{T-1}$ for each self-employment choice. With $\bar{K}_{T-1}$ in hand, I then estimate the choice-specific value function as discussed above $\bar{V}_T(d_{t-1}) = (E[V_T|d_{T-1}, Z_{T-1}])$ for each choice at time $T-1$. This process can then be iterated back to time $t = 1$. Specifically, the likelihood contribution at time $t < T$ is given by:

$$\mathcal{L}_t^K = \int \prod_i \prod_{\delta \in D} \left(f_\delta(K_t | \boldsymbol{Z}_t, P(d_{t+1} = \delta | \boldsymbol{Z}_t, \eta), s, \eta)\right)^{\mathbb{1}\{d_t = \delta\}} f(\eta) d\eta \tag{24}$$

$$\mathcal{L}_t^Y = \int \prod_i \prod_{\delta \in D} \left(f_\delta(Y_t | \boldsymbol{Z}_t, \bar{K}_{d_t}, s, \eta)\right)^{\mathbb{1}\{d_t = \delta\}} f(\eta) d\eta \tag{25}$$

$$\mathcal{L}_t^d = \int \prod_i f(d_t | \boldsymbol{Z}_t, \bar{\boldsymbol{Y}}_t, s, \bar{v}_{t+1}(d_t), \eta) f(\eta) d\eta \tag{26}$$

Finally, the choice of schooling contributes to the likelihood:

$$\mathcal{L}^s = \int \prod_i f(s_i | \boldsymbol{Z}_0, \bar{v}_1(s), \eta) f(\eta) d\eta \tag{27}$$

The overall likelihood can then be written as:

$$\mathcal{L} = \mathcal{L}^s \times \prod_{t=1}^{T} \left(\mathcal{L}_t^K \mathcal{L}_t^Y \mathcal{L}_t^d\right). \tag{28}$$

G Additional Results on Goodness of Fit

This section provides additional results documenting the goodness of fit of the model. Table 22 shows the proportion of individuals in each career by potential experience for both the simulation and the data. Table 23 shows results from regressing income on predicted income from the model. The first column shows this regression for thee full population, while the other columns show the regression conditional on career.

Table 22: Proportion in Each Career by Potential Experience (simulation and data)

	Simulation						
Pot. Exp.	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC
0	0.20	0.27	0.50	0.00	0.01	0.01	0.01
1	0.20	0.27	0.50	0.00	0.01	0.01	0.01
2	0.25	0.27	0.45	0.00	0.01	0.01	0.01
3	0.24	0.29	0.42	0.01	0.01	0.01	0.01
4	0.21	0.27	0.47	0.01	0.01	0.01	0.02
5	0.18	0.28	0.49	0.01	0.01	0.01	0.02
6	0.18	0.28	0.48	0.01	0.02	0.01	0.02
7	0.16	0.29	0.48	0.01	0.02	0.01	0.03
8	0.14	0.30	0.49	0.01	0.02	0.02	0.03
9	0.11	0.32	0.49	0.01	0.02	0.02	0.03
10	0.10	0.33	0.49	0.01	0.02	0.02	0.03
11	0.09	0.33	0.49	0.02	0.03	0.02	0.03
12	0.09	0.33	0.49	0.02	0.03	0.02	0.03
13	0.09	0.33	0.48	0.02	0.03	0.02	0.03
14	0.09	0.32	0.48	0.02	0.03	0.02	0.04
15	0.09	0.32	0.49	0.02	0.03	0.02	0.04
16	0.08	0.32	0.49	0.02	0.03	0.02	0.04
17	0.07	0.32	0.50	0.02	0.04	0.02	0.04
18	0.07	0.32	0.49	0.02	0.04	0.02	0.04
19	0.08	0.32	0.48	0.02	0.04	0.02	0.04
20	0.08	0.32	0.48	0.03	0.04	0.02	0.04
21	0.07	0.32	0.49	0.03	0.04	0.02	0.03
22	0.05	0.33	0.50	0.03	0.05	0.02	0.03
23	0.05	0.32	0.51	0.03	0.05	0.01	0.03
	Data						
Pot. Exp.	Non-Emp	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC
0	0.23	0.28	0.48	0.00	0.00	0.00	0.01
1	0.19	0.31	0.49	0.00	0.01	0.00	0.01
2	0.21	0.32	0.45	0.00	0.01	0.00	0.01
3	0.22	0.32	0.43	0.00	0.01	0.00	0.01
4	0.17	0.34	0.46	0.01	0.01	0.01	0.01
5	0.14	0.34	0.48	0.01	0.01	0.01	0.01
6	0.14	0.34	0.48	0.01	0.01	0.01	0.02
7	0.12	0.34	0.48	0.01	0.02	0.01	0.02
8	0.11	0.34	0.49	0.01	0.02	0.01	0.02
9	0.10	0.34	0.49	0.01	0.02	0.01	0.02
10	0.09	0.34	0.50	0.02	0.02	0.01	0.03
11	0.08	0.33	0.50	0.02	0.03	0.01	0.03
12	0.08	0.32	0.50	0.02	0.03	0.01	0.03
13	0.08	0.32	0.51	0.02	0.03	0.01	0.03
14	0.08	0.31	0.51	0.02	0.04	0.02	0.03
15	0.08	0.29	0.52	0.02	0.04	0.02	0.03
16	0.08	0.28	0.53	0.02	0.04	0.02	0.04
17	0.07	0.27	0.54	0.02	0.04	0.02	0.04
18	0.07	0.26	0.55	0.02	0.05	0.01	0.04
19	0.09	0.25	0.54	0.02	0.05	0.01	0.04
20	0.08	0.24	0.55	0.02	0.05	0.01	0.04
21	0.07	0.23	0.56	0.02	0.06	0.01	0.04
22	0.05	0.24	0.57	0.02	0.06	0.01	0.04
23	0.06	0.24	0.57	0.02	0.06	0.01	0.04

Notes: Table shows the proportion of individuals in each career by potential experience in the simulated model and in the data. "WC" stands for white-collar, "BC" stands for blue-collar, "PE" stands for paid employment. "SE-I" stands for incorporated self-employment. "SE-U" stands for unincorporated self-employment.

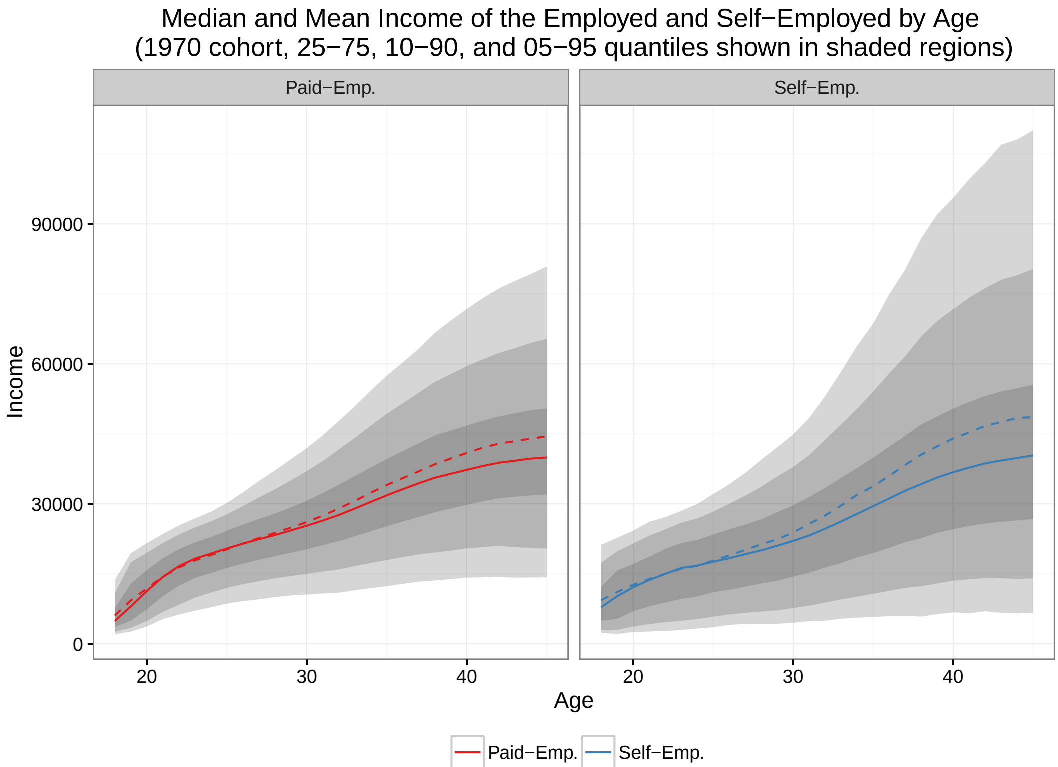
Table 23: Regressing Log wage Income on Predicted Log Wage Income.

	Log Income						
	ALL	PE WC	PE BC	SE-I WC	SE-I BC	SE-U WC	SE-U BC
Pred. Log Income	0.989***	1.012***	1.000***	0.994***	0.991***	0.988***	0.999***
	(0.001)	(0.001)	(0.001)	(0.009)	(0.006)	(0.015)	(0.009)
Constant	0.038***	−0.046***	−0.002	0.023	0.037	0.036	0.001
	(0.002)	(0.005)	(0.004)	(0.035)	(0.022)	(0.046)	(0.028)
R^2	0.692	0.694	0.694	0.542	0.562	0.369	0.373

Notes: Table shows the regression of log income on predicted log income from the structural model. The first column shows the regression across the full population why the remaining columns show these regressions conditional on career. *p<0.1; **p<0.05; ***p<0.01

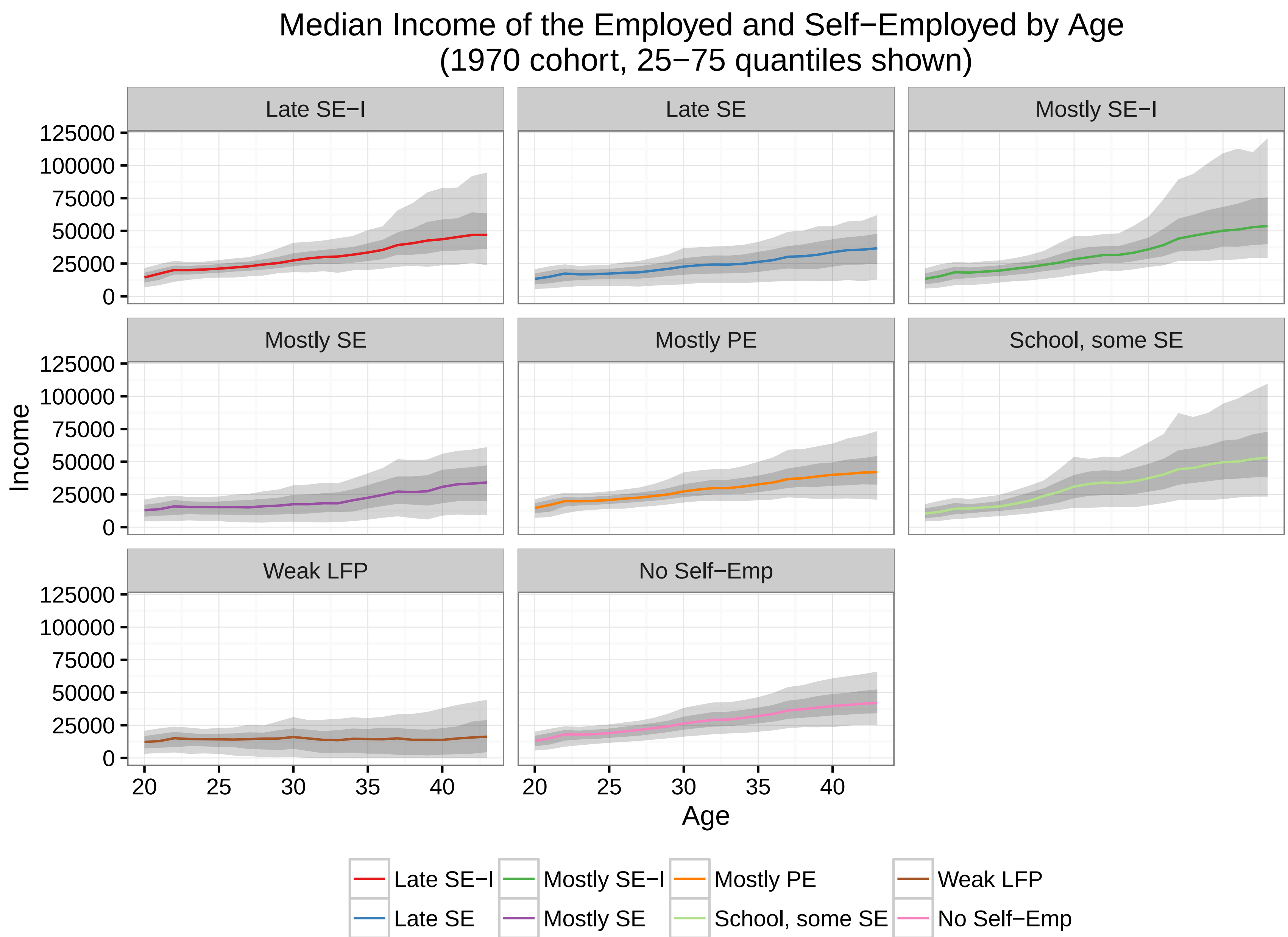
H Additional Results

Figure 11: Earning of the Employed and Self-Employed over the life cycle



Notes: Figure shows the mean and median after-tax earnings of the self-employed and employed by age. The shaded region shows the 25-75 quantile range. All amounts are in 2010 U.S. dollars

Figure 12: Earning of the Employed and Self-Employed over the life cycle. (by cluster)



Notes: Figure shows the mean and median after-tax earnings by career trajectory cluster. Shaded regions show the 25-75, 10-90, and 05-95 quantile ranges. All amounts are in 2010 U.S. dollars