Self-employment dynamics & the importance of the state-dependent effect

Darcy Fitzpatrick

National Institute of Labour Studies – Flinders University, South Australia

JEL Classifications: C33, J21, J24, J62, L26

Keywords: Self-employment, entrepreneurship, labour market dynamics, dynamic random-effects panel probit models, state-dependence, unobserved heterogeneity, Household Income & Labour Dynamics in Australia (HILDA) survey

Acknowledgements: This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the author and should not be attributed to either DSS or the Melbourne Institute.

1 Abstract

(150 words)

2 Introduction

The role of self-employment in the labour market and the factors that influence workers to choose self-employment has been extensively studied in the economic literature. However, much of the empirical evidence on this topic remains incoherent and inconclusive. On the one hand, self-employment is often considered as a form of entrepreneurship.[1] This interpretation regards self-employment as a preferred or desired outcome because of the increased pecuniary and non-pecuniary rewards the opportunity offers to individuals, as well as being encouraged by governments and policy-makers as a means of promoting economic growth and

[1] See Shane & Venkataraman (2000), *inter alia*.

employment generation.[2] On the other hand, self-employment is considered from a more pessimistic viewpoint as either: a form of 'disguised unemployment', which attracts displaced and redundant workers due to business-cycle fluctuations or structural changes in the economy[3]; or, as an occupation of 'last-resort', which attracts workers of 'poor' quality — such as, those with histories of weak attachment to the labour market, poor skills or low ability, or those suffering from racial, gender, or age discrimination by employers.[4] Another possibility is that self-employment encompasses both positive and negative viewpoints depending on the types of worker and the prevailing economic circumstance.

It is also no coincidence that much of the existing research is based predominately on cross-sectional (or pooled panel) data methods of analysis. Across the literature, the prevalence of these models have resulted in the identification of a vast number of statistically significant relationships that are often inaccurate and contradictory to one another, contributing a 'scattershot' or obscured understanding of self-employment. Instead, a clearer understanding of self-employment and why it is chosen by certain individuals can be gained by utilising longitudinal data and econometrically advanced panel data techniques.

Panel data models observe the changes in the behaviour and responses of individuals over an extended period of time, and estimate with an improved robustness sources of observed and unobserved heterogeneity that are otherwise neglected in cross-sectional models. A further advantage of panel models is the ability to include dynamic extensions and control for the possible influence of 'genuine' state-dependence: that is, the extent to which individuals' past experience or 'lagged' self-employment status in a previous period *in and of itself* affects the current self-employment status.

In other areas of labour economics, the use of dynamic panel techniques to account for the possibility of unobserved heterogeneity and state-dependence has proved insightful for isolating the causal relationships that determine, for example,

[2] See Reynolds et al. (2004), Rotefoss & Kolvereid (2005), Parker & Belghitar (2006), Henley (2007), van Praag & Versloot (2007) and Stam et al. (2008), *inter alia*, for a discussion of the entrepreneurial reasons for self-employment participation.
[3] See Covick (1998), Earle & Sakova (2000), Andersson & Wadensjö (2007), Hyytinen & Rouvinen (2008) and von Greiff (2009), *inter alia*, for a discussion of self-employment as a form of disguised unemployment or because of displacement from salaried-employment.
[4] See Evans & Leighton (1989), Clark & Drinkwater (2000) and Parker & Rougier (2007), *inter alia*, for a discussion of self-employment as an occupation of last resort for certain workers.

unemployment and low-pay employment states.[5] By contrast, only a few studies (predominately European) analyse the importance dynamics in determining self-employment outcomes.[6] These studies find strong and consistent evidence that self-employment is a persistent state and that this is largely a product of state-dependence. What is not clear from these studies, however, is the extent to which controlling for unobserved heterogeneity and state-dependence change the story about self-employment and help elucidate the contrapuntal themes identified in the cross-sectional research. This is the purpose of this study.

Using the first eleven waves of the HILDA survey for the period 2001 to 2011, this study models the self-employment status of Australian workers using both static cross-sectional and dynamic panel data techniques to analyse the sensitivity of the results to changes in the method of estimation. This approach is most similar to Henley's (2004) analysis of self-employment dynamics and the influence of state-dependence on self-employment in the U.K., and there appear to be no other comparable Australian studies. This study finds the change in the results obtained between the cross-sectional and the dynamic panel models to be considerable. Most of the observed characteristics are rendered either statistically or economically insignificant once dynamics are adequately accounted for. This is in stark contrast to the importance placed on the observed characteristics in much of the, predominately cross-sectional, evidence on self-employment and casts doubt on the validity of the findings in the existing research.

Moreover, this study confirms the findings from previous dynamic panel studies of self-employment that genuine state-dependence is an important influence on individuals' self-employment status. That is, while certain individual characteristics and traits, particularly those that are persistent and unobservable (e.g. cognitive and non-cognitive abilities), exacerbate the likelihood of self-employment, it is the influence of past experience *itself* that leaves workers prone to continued participation in self-employment. This result also demonstrates the importance of addressing the potential for endogeneity in the initial condition. However, despite establishing the importance of genuine state-dependence of self-employment, the possible causes underlying the state-dependence remain unresolved.

[5] See Heckman (1981a), Stewart & Swaffield (1999), Uhlendorff (2006), Stewart (2007), Cappellari & Jenkins (2008), Buddelmeyer et al. (2010), Cai (2014), *inter alia*.
[6] See Henley (2004) for the U.K., Caliendo & Uhlendorff (2008) for Germany, and Taylor (2011) for selected E.U. countries.

3 Previous Research

The determinants of self-employment and the reasons why people become self-employed is well-documented in the existing economic literature (see Le (1999), Sørensen & Chang (2006) Praag & Versloot (2007), Sluis et al. (2008), and Parker (2009: 106-157) for overviews of this literature). The salient determinants of self-employment that emerge in the existing research include: a wage differential between self-employment and salaried-employment, the intergenerational transfer of self-employment, wealth accumulation and access to finance, spousal support, education attainment and the acquisition of skills through experience, regional differences or 'thin' labour markets, structural economic changes and business-cycle fluctuations, risk aversion and tolerance of uncertainty, and other personality traits thought to favour entrepreneurial behaviour. However, the usefulness of the findings from much of this research is questionable. A large number of individual characteristics are identified throughout the research as having a statistically significant relationship with self-employment status, but with very little consistency. As summarised by Parker (2009), the estimates from some 153 scholarly articles generate an exhaustive list of significant determinants, as well as a considerable amount of contradiction in the direction of the estimated effects (positive/zero/negative) – reproduced in Table 1.

Table 1: Summary of significant determinants of self-employment (Parker, 2009)

Explanatory variable	No. +	No. –	No. 0
Income differential	8	2	4
Age	83	6	14
Experience	24	1	2
Education	69	21	27
Risk aversion	0	11	3
Married / working spouse	52	9	8
Number of children	16	2	3
Ill health / disability	5	4	6
Entrepreneur parent	40	2	2
Technological progress	4	4	2
Unemployment			
Cross-section	22	14	18
Time series	33	5	2
Urban location	7	7	4
Immigration	5	1	0
Interest rates	1	9	3
Personal wealth	40	2	4
Personal income tax rates	12	5	1

Notes: +, − and 0 denote positive (significant at a 5% level), negative (significant at a 5% level), and zero (insignificant at 5% level) coefficients, respectively. Only multivariate studies (i.e. those including controls for other explanatory variables) are included; descriptive studies are excluded. For row 11, panel studies with large *N* and small *T* are classified as cross-section; those with large *T* and small *N* are classified as time-series.

Source: Reproduced from Parker (2009: 108)

A potential explanation for the lack of consensus in the existing empirical research is its heavy reliance upon cross-sectional (or pooled panel) data to estimate static models of self-employment.[7] Such models estimate the dependence of workers' probability of being self-employed purely on differences in the observed characteristics between those in self-employment and salaried-employment, at a point in time. However, the effectiveness of these models at capturing the self-employment choice is most unsatisfactory. Because static cross-sectional models neglect to control for the potential influence of unobserved heterogeneity, the estimates are likely to overstate the importance of the observed characteristics and suffer from omitted variable bias. A further limitation of the static cross-sectional approach is its assumption that workers face both the self-employment and salaried-employment opportunities concurrently. Because the current self-employment choice is determined independent of the self-employment status in the past, the estimates confound both the determinants of transition to self-employment with those of survival in self-employment (Evans & Leighton, 1989).

Historically, the prevalence of studies using static cross-sectional models has been due to the absence for many decades of longitudinal surveys in many developed

[7] See Rees & Shah (1986), Evans & Leighton (1989), Taylor (1996), Blanchflower & Oswald (1998), Dunn & Holtz-Eakin (2000), Hamilton (2000) and Clark & Drinkwater (2000), *inter alia*.

countries, limiting the availability of more adequate panel data estimation techniques to researchers. However, even as the availability of panel data and associated econometric techniques has increased, the number of studies that fully exploit panel data to analyse the dynamics of self-employment is far less developed than in other areas of labour economics research. In the self-employment literature, only a handful of studies analyse the individual self-employment participation decision in a dynamic panel framework (see, for example, Henley, 2004; Caliendo & Uhlendorff, 2008; Taylor, 2011). These studies model the dynamics of transitions between self-employment and salaried employment over time, and identify the extent to which the persistence of self-employment is caused by genuine state-dependence and observed/unobserved individual heterogeneity. Unlike the inconsistent findings of the earlier studies, the dynamic panel studies find strong and consistent evidence that state-dependence is an important determinant of self-employment, both statistically and economically, even once observed and unobserved heterogeneity is accounted for.

Using the first 9 waves of the British Household Panel Survey (BHPS), Henley (2004) finds that relative to those working in wage-jobs, the workers who were self-employed in the previous year increased their probability of being self-employed in the current year by approximately 30 percentage points, and that the unobserved individual heterogeneity accounts for approximately 60% of the unexplained variance of the composite error. Consistent with these findings, Taylor (2011) obtains similarly large estimates, between 20 to 89 percentage points, for male workers from selected European countries (using the European Community Household Panel (ECHP), 1994-2011), as does Caliendo & Uhlendorff (2008), 22 percentage points, for German workers (using the Socio-Economic Panel (SOEP), 1984-2005). These studies also find substantial correlation between the unobserved heterogeneity and the initial condition, highlighting the importance of treating the initial conditions as endogenous.

While the existing dynamic panel studies distinguish the importance of genuine state-dependence from the influence of heterogeneity on self-employment status, no meaningful attempt is made to distinguish the possible mechanisms causing the state-dependence. Unlike unemployment, for example, where the reasons for state-dependence are obvious because of its undesirable nature and the negative

consequences it has on individual's wellbeing, such conclusions are difficult to draw when considering self-employment against salaried-employment. Past experience in self-employment may be interpreted as having either a 'scarring' or 'virtuous' effect on the current self-employment outcome depending on how self-employment and its outcomes are perceived.

There is contradictory descriptive evidence that perceive self-employment as having both positive and negative outcomes. On the one hand, self-employment is often optimistically perceived and interpreted as a form of entrepreneurship. For example, as reported by Blanchflower et al. (2001), in most developed economies a large minority share of employees hold a latent or unfilled desire to instead work in self-employment. Similarly, most workers who make the transition into self-employment are employees, rather than unemployed (Evans & Leighton, 1989; Henley, 2004), and who appear to make the transition voluntarily (Farber, 1999). Self-employment is also associated with providing workers with greater non-pecuniary rewards, expressed as higher levels of work satisfaction by those in self-employment (Blanchflower & Oswald, 1998; Blanchflower, 2000; Hundley, 2001; Benz & Frey, 2008b; Benz & Frey, 2008a). Based on this evidence, state-dependence of self-employment might be interpreted as having a 'virtuous' effect where self-employment is less accessible than ostensibly thought and employees are 'locked-out' from transitioning to self-employment.

On the other hand, there is a substantial amount of descriptive evidence that contrasts with the optimistic perception of self-employment. In Australia, a large portion of the work undertaken in self-employment includes many mundane trades and professional occupations — such as, livestock and crop farming, truck driving, plumbing, electrician, hairdressing, and bookkeeping — which would not typically be considered as entrepreneurial activities.[8] Furthermore, much of this self-employment activity does not create employment growth or generate additional labour demand, as most self-employed are own-account workers, that is, they work alone.[9] Also, for many workers, self-employment has worse employment outcomes than they could otherwise expect working as an employee. For example, self-employed workers are observed to experience both longer-working hours

[8] Based on detailed occupation statistics from the ABS 2011 Census of Population & Housing
[9] Own-account workers accounted for approximately 63% of total self-employment in Australia based on statistics from the ABS 2008 LFS-FoE

(Hyytinen & Ruuskanen, 2007) as well as suffering a wage-penalty (Hamilton, 2000). There is also evidence that self-employed workers experience higher levels of stress and anxiety, which has a negative effect on the health and familial situations of workers (Blanchflower, 2004; Taris et al., 2008; Stephan & Roesler, 2010). Based on this evidence, the effect of past experience in self-employment may be interpreted as having a 'scarring' effect that generates conditions that trap self-employed workers and reduce their chance of escaping to salaried-employment in the future

4 Data sources & defining self-employment

Throughout this study, two data sources are used to analyse the dynamics of self-employment. The principal source of data, used for the multivariate analysis segment of this study, is the first eleven waves of the HILDA longitudinal survey for the period 2001 to 2011. The HILDA survey and its use in this study are very similar to the sources of longitudinal data used to perform similar multivariate analysis in the existing dynamic panel studies of self-employment.

Detailed descriptions of HILDA, its history and its uses are well documented.[10] The HILDA survey is a representative sample survey of the Australian population that, since 2001, has attempted to follow the same individuals on an annual basis. The HILDA also collects a breadth of detailed information on a range of topics, including: household and familial relationships and background, demographic characteristics, education and training issues, labour market experience and employment arrangements, income and expenditure, time-use, social and lifestyle issues, and health and well-being. For the period 2001 to 2011, the HILDA survey had an unbalanced sample of 26,028 Australian residents, aged 15 years and over, from across 7,682 households. During this period, the characteristics of the responding sample remained a relatively good-match to the Australian population at a broad level (Watson & Wooden, 2013).

To complement HILDA, this study also utilises additional data from the 2006, 2010 and 2012 ABS LFS-LM cross-sectional surveys as auxiliary sources of information for comparative and descriptive purposes. The LFS-LM surveys are representative sample surveys of the Australian population and labour force that provide a

[10] See Wooden et al.(2002), Watson & Wooden (2002), Wooden & Watson (2007), Watson & Wooden (2012) and Richardson (2013), *inter alia*.

biennial snapshot of the employment arrangements of workers and the timing and duration of transitions events for the preceding 12 months. In comparison to HILDA, the LFS-LM surveys are much larger and typically collect information on approximately 32,000 respondents from 28,000 households. However, unlike HILDA, the LFS-LM surveys are narrower in breadth with respect to the information collected. The cross-sectional nature of the LFS-LM surveys also mean that the statistics potentially suffer from 'recall bias' through respondents intentionally misrepresenting or unintentionally mis-recollecting events that transpired during the preceding 12 months.

While there is considerable variation in the definition and enumeration of self-employment in the existing research, this study considers self-employment to be a mutually exclusive labour market state that is distinct from salaried-employment, unemployment, and not-in-the labour force. In keeping with the conventional approach to defining self-employment in labour economics, self-employment is broadly defined as: those whose 'remuneration is directly dependent upon profits and the incumbents make operational decisions, or delegate such decisions, while retaining responsibility for the welfare of the enterprise'.[11] Fortunately, both the HILDA and LFS-LM surveys are similar in their classification of labour force status and employment arrangements of workers and both allow for the accurate identification of self-employed workers.

From the data, the classification of self-employment refers to the aggregation of: *owner-managers of an unincorporated enterprise* (OMUE)[12]; *owner-managers of an incorporated enterprise* (OMIE)[13]; and, *contributing-family workers*[14]. This classification of self-employment also includes *own-account workers* (i.e. those who work alone) and *employers* (i.e. those who employ additional labour). In contrast to much of the existing economic research related to self-employment, the classification of self-employment in this study explicitly includes OMIEs. The inclusion of OMIEs in self-employment is substantial and accounts for approximately 35% of the sample of self-employed workers in the HILDA data, and 35% and 38% of self-employment in

[11] International Classification of Status in Employment (ICSE-93)
[12] This is where the worker is remunerated directly from the profits of their business, but there is also no legal distinction between the personal liabilities of the worker and the assets of their business.
[13] This is where the worker and their business are separate legal entities and the worker is employed under the account of the business (a limited liability company), but the worker retains a controlling interest and remains responsible for its operation and is entitled to a distribution of the profits.
[14] This is where the worker works in a family owned and operated business and without explicitly being paid, but may benefit implicitly from the proceeds of the business.

the 2006 and 2012 samples of the LFS-LM survey, respectively.

5 Aggregate labour mobility of self-employment in Australia

Self-employment in Australia is a highly persistent labour market state for most workers. Table 2 summarises the aggregate year-on-year transitions between each of the labour market states for the years 2006 and 2012 using the LFS-LM data and for the pooled 2002-2011 period using the HILDA data. In 2012, the share of self-employed workers who had previously been self-employed in 2011 ($t-1$) account for 91.7% of the self-employed workforce. The high incidence of persistence in self-employment is also comparable to the share of employees who remained in salaried-employed jobs year-on-year (89.2%, in 2012), while dissimilar to those who remained unemployed or not-in-the labour force (26.5% and 23.4%, respectively, in 2012). The patterns of persistence and mobility, particularly in self-employment, also remain relatively consistent over time between 2006 and 2012.

As discussed earlier, differences in the collection methodologies between the LFS-LM and HILDA surveys also produce some striking statistical differences. As evident from the statistics in Table 2, in comparison to the LFS-LM data, the HILDA data captures a similar pattern of year-on-year transitions for salaried-employment, while capturing a much higher incidence of mobility for self-employment. For example, in the HILDA data, only 80.8% of workers persist in self-employment year-on-year in comparison to 91.7% in the 2012 LFS-LM data. The discrepancies in the estimates between the HILDA and the LFS-LM surveys cannot be fully explained. However, in the context of this study, the over-enumeration of the transitions involving self-employment in the HILDA data is seen as beneficial.

Table 2: Mobility of labour & the transitions between labour market states, $t-1$ to t

Labour market transitions		ABS LFS-LM survey 2006		ABS LFS-LM survey 2012		HILDA survey 2002-2011	
t-1	*t*	N	*%*	N	*%*	Obs.	*%*
Employee	Employee	23,377	*88.6*	21,684	*89.2*	53,388	*90.5*
Self-employed	Employee	176	*0.7*	116	*0.5*	1,471	*2.5*
Unemployed	Employee	-	-	-	-	1,476	*2.5*
NILF	Employee	-	-	-	-	2,638	*4.5*
Unemp./NILF	Employee	2,832	*10.7*	2,496	*10.3*	-	-
	Total	26,385	*100.0*	24,296	*100.0*	58,973	*100.0*
Employee	Self-employed	323	*5.0*	188	*3.5*	1,692	*13.6*
Self-employed	Self-employed	5,779	*89.7*	4,899	*91.7*	10,060	*80.8*
Unemployed	Self-employed	-	-	-	-	96	*0.8*
NILF	Self-employed	-	-	-	-	609	*4.9*
Unemp./NILF	Self-employed	337	*5.2*	257	*4.8*	-	-
	Total	6,439	*100.0*	5,344	*100.0*	12,457	*100.0*
Employee	Unemployed	637	*66.8*	553	*67.1*	1,084	*38.2*
Self-employed	Unemployed	44	*4.6*	53	*6.4*	105	*3.7*
Unemployed	Unemployed	-	-	-	-	824	*29.0*
NILF	Unemployed	-	-	-	-	824	*29.0*
Unemp./NILF	Unemployed	273	*28.6*	218	*26.5*	-	-
	Total	954	*100.0*	824	*100.0*	2,837	*100.0*
Employee	NILF	1,184	*63.7*	1,095	*65.7*	3,199	*8.9*
Self-employed	NILF	189	*10.2*	182	*10.9*	778	*2.2*
Unemployed	NILF	-	-	-	-	807	*2.2*
NILF	NILF	-	-	-	-	31,354	*86.8*
Unemp./NILF	NILF	486	*26.1*	390	*23.4*	-	-
	Total	1,859	*100.0*	1,667	*100.0*	36,138	*100.0*
	Total (N)	35,637	-	32,131	-	110,405	-

Notes: Unweighted estimates. HILDA survey data is pooled and unbalanced. Unemp. = unemployed; NILF = not-in-the labour force

Source: ABS, cat. no. 6202.0.30.004, Labour Force Survey and Labour Mobility, Australia, Feb 2012, CURF (Expanded) – accessed via RADL 20140223
ABS, cat. no. 6202.0.30.004, Labour Force Survey and Labour Mobility, Australia, Feb 2006, CURF (Basic)
HILDA Survey, 2001-2012

The high incidence of persistence in self-employment provides no insight about the labour market pathways where self-employed workers arrive from or depart to. To better gauge the direction and relevance of the inflow and outflow transitions between self-employment and the other labour market states, Table 3 describes the aggregate inflow and outflow transitions involving self-employment relative to those who remained in self-employment. Much of the labour mobility involving self-employment occurs from within the labour market, predominately by employees already actively engaged in salaried-employed jobs. The incidence of transitions between self-employment and salaried-employment, in comparison to the unemployed and not-in-the labour market states, is particularly acute in the pooled 2002-2011 HILDA data more so than in the 2006 and 2012 ABS LFS-LM data. For example, relative to those who remain in self-employed, the HILDA data estimates the year-on-year inflow and outflow transitions between self-employment

and salaried-employment to be 16.8% and 14.6%, respectively; whereas, the 2012 LFS-LM data estimates the same relative inflow and outflow transitions to be 3.8% and 2.4%, respectively. By comparison, very few of the transition pathways to and from self-employment involve unemployment. For example, in the HILDA data, the total number of year-on-year transitions between self-employment and unemployment only accounted for 2.0% of the persistent self-employed workforce.

Moreover, the direction of the transitions to and from self-employment is slightly unbalanced, as the relative shares of workers entering self-employment are slightly greater than the amount exiting from self-employment. This imbalance in the flows is also further pronounced when the transition pathways for each state are considered separately. It appears that, in aggregate, employees who enter self-employment from salaried-employment are more likely to exit self-employment to non-employment (most likely exiting the labour market completely), while workers entering from non-employment are less likely to exit from self-employment to salaried-employment. However, this provides no indication about the permanency of self-employment over the life-cycle of workers.

Table 3: Relative inflow & outflow transitions between self-employment & other labour market states, $t-1$ to t

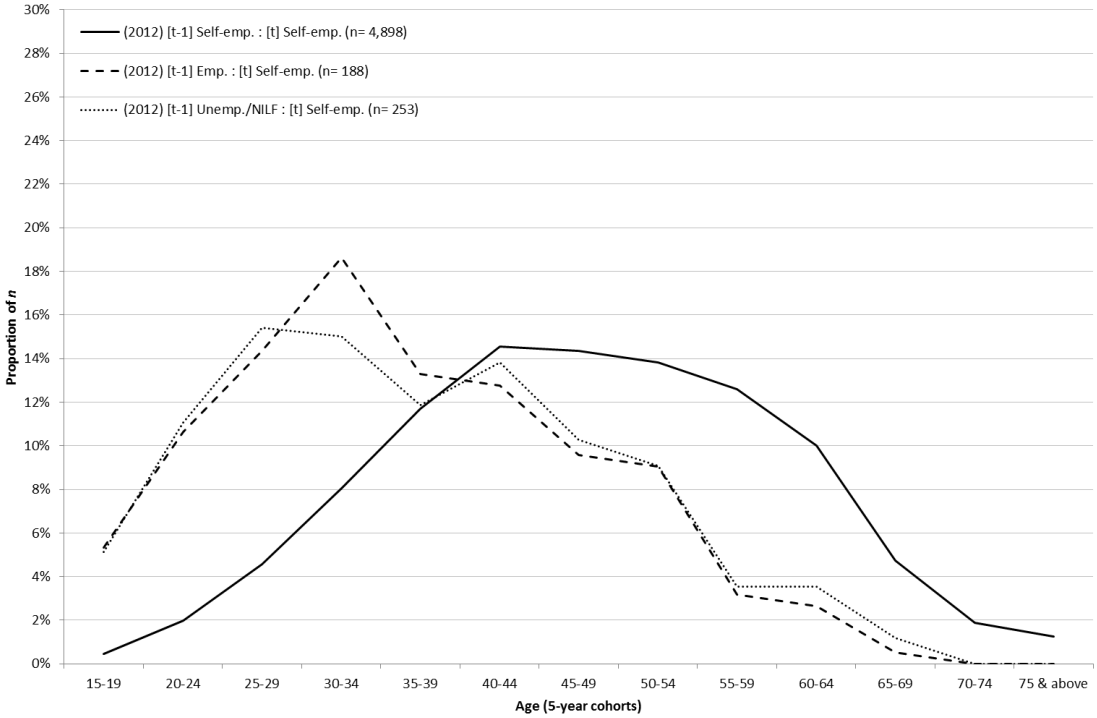
		ABS LFS-LM survey				HILDA survey	
		2006		2012		2002-2011	
		N	*Transition / Stayers (%)*	N	*Transition / Stayers (%)*	Obs.	*Transition / Stayers (%)*
Stayers:							
t-1	*T*						
Self-employed	Self-employed	5,779	*100.0*	4,899	*100.0*	10,060	*100.0*
Inflow transitions:							
t-1	*T*						
Employee	Self-employed	323	*5.6*	188	*3.8*	1,692	*16.8*
Unemployed	Self-employed	-	-	-	-	96	*1.0*
NILF	Self-employed	-	-	-	-	609	*6.1*
Unemp./NILF	Self-employed	337	*5.8*	257	*5.2*	705	*7.0*
Outflow transitions:							
t-1	*T*						
Self-employed	Employee	176	*3.0*	116	*2.4*	1,471	*14.6*
Self-employed	Unemployed	-	-	-	-	105	*1.0*
Self-employed	NILF	-	-	-	-	778	*7.7*
Self-employed	Unemp./NILF	233	*4.0*	235	*4.8*	883	*8.8*

Notes: see Table 2
Source: see Table 2

A clearer picture about the sequence or order in which workers engage in self-employment is gained by describing the timing of the inflow and outflow transitions over the life-cycle. Comparing the age-distributions of the inflow

transitions to self-employment and salaried-employment, illustrated in Figure 1 and Figure 2 respectively, the opportunities in self-employment appear to be relatively more attractive to older workers who were already actively engaged within the labour market with skills and experience. As shown in Figure 1, workers enter self-employment from both the salaried-employment and non-employment states at a similarly older age, at a mid-point in the work-life cycle (approx. 30-34 years). By contrast, the inflow transitions to salaried-employment (Figure 2) are predominately young and inexperienced workers who enter from outside the labour market or from unemployment at the initial or early stages of the work-life cycle (approx. 15-29 years). Furthermore, once engaged, the evidence suggests that self-employment is permanent form of employment for most self-employed over the work-life cycle. The age distribution of the workers entering self-employment is predominately younger than those who exit. While the self-employed who do not survive predominately exit to salaried-employment sooner rather than later (approx. 35-39 years, as shown in Figure 2), the majority of the outflow transitions from self-employment to non-employment occur over an older, more elongated age range (approx. 45-66 years, as shown in Figure 3). Self-employed workers are also more likely continue to work past the conventional age of retirement for employees (approx. 55-65 years), remaining in self-employment until their 70's (as shown in Figure 3).

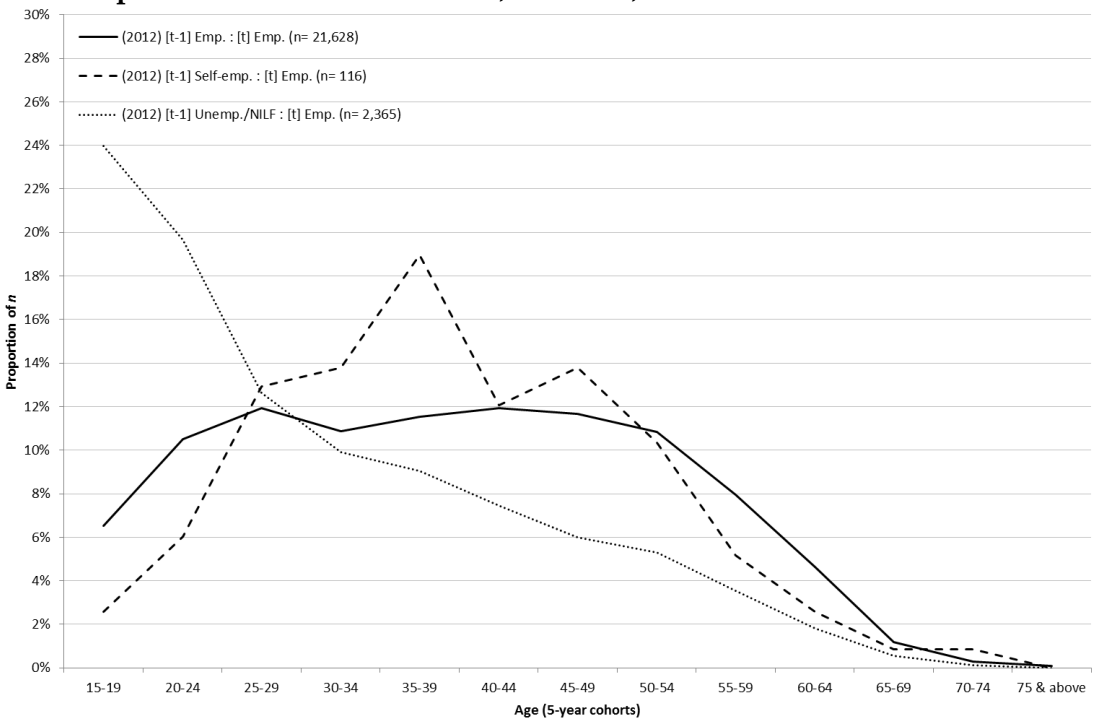
Figure 1: Age distributions of inflow transitions to self-employment by previous labour market state, $t-1$ to t, in 2012



Notes: Unweighted estimates. Emp. = Employee; Self-emp. = Self-employed; Unemp. = unemployed; NILF = not-in-the labour force

Source: ABS, cat. no. 6202.0.30.004, Labour Force Survey and Labour Mobility, Australia, Feb 2012, CURF (Expanded) – accessed via RADL 23/02/2014

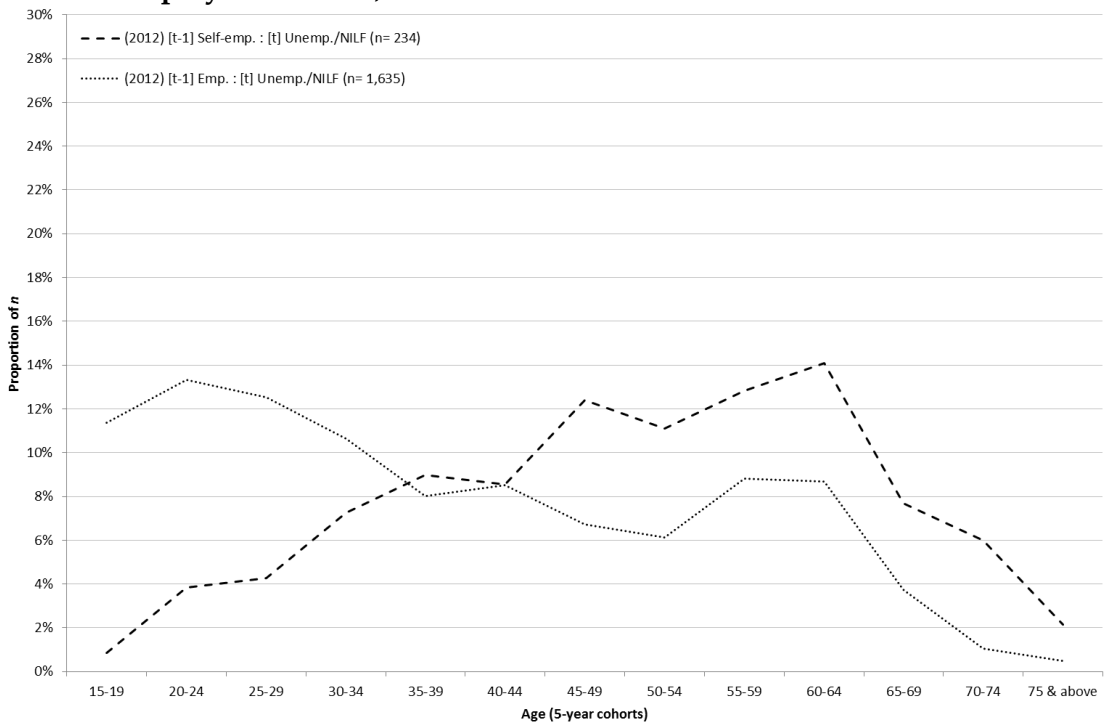
Figure 2 Age distributions of inflow transitions to salaried-employment by previous labour market state, $t-1$ to t, in 2012



Notes: see Figure 1

Source: see Figure 1

Figure 3 Age distributions of outflow transitions from employment to non-employment, $t-1$ to t, by previous self-employment and salaried-employment status, in 2012



Notes: see Figure 1
Source: see Figure 1

There is also evidence to suggest that the participation in self-employment occurs as part of a career choice or progression, rather than behaviour that is forced by circumstance or necessity. Table 4 presents the proportion of transitions that were reported as voluntarily or involuntarily for those who ceased working in either a salaried-employed job or self-employment in year $t-1$. In 2012, for example, 76.6% of employees who transitioned to self-employment cited a voluntary reason for the transition — such as, unsatisfactory work conditions, better job opportunity, to start a new business, or family reasons. The share of employees voluntarily transitioning to self-employment is also greater than the share of employees who changed salaried-employed jobs (72.3%, in 2012).

On the whole, the descriptive analysis appears to indicate that self-employment is a destination *in itself*, which for many workers is a desired outcome but only ever achieved by a few. However, drawing inferences from the transition probabilities and persistence of self-employment in aggregate is difficult. The propensity for certain workers to transition and persist in self-employment does not necessarily imply that the state-dependence observed in aggregate is true for individuals, because there is more than one possible explanation for persistence (Heckman,

1981a). At the individual level, the probability of participating in self-employment could also be the result of persistent but unobserved individual characteristics. Therefore, additional controls for both observed past behaviour and unobserved characteristics are necessary for more accurate inferences about self-employment.

Table 4: Reason for ceasing job (involuntary/voluntary) and transitioning between labour market states, $t-1$ to t

Labour market transitions		ABS LFS-LM survey 2012			
			Invol.	Vol.	Non-resp.
t-1	*t*	N	*% (row)*	*% (row)*	*% (row)*
Employee	Emp. (new job)	2,655	*21.7*	*72.3*	*6.0*
Self-employed	Employee	116	*14.7*	*77.6*	*7.8*
Employee	Self-employed	188	*20.2*	*76.6*	*3.2*
Employee	Unemp./NILF	1,648	*42.5*	*57.5*	*0.0*
Self-employed	Unemp./NILF	235	*40.0*	*60.0*	*0.0*

Notes: Unweighted estimates. Emp. = Employee; Unemp. = Unemployed; NILF = not-in-the labour force; Invol. = involuntary; Vol. = voluntary; Non-resp. = missing information.
Aggregation of reasons for transition between labour market states was defined by ABS (2013).
Source: see Table 2

6 Multivariate analysis of self-employment dynamics

6.1 Econometric model

The salient point of this study considers whether individuals' choice of self-employment is a product of their past observed and unobserved behaviours, and to what extent controlling for the dynamics of self-employment alters how self-employment is understood. In doing so, this study models the self-employment status of Australian workers using cross-sectional and panel data techniques.

In keeping with much of the existing research, the sequence begins by modelling the probability of self-employment in a static cross-sectional (or pooled-panel) probit framework, where the current outcome is determined by differences in the distribution of observed individual heterogeneity of the workers in self-employment relative to employees in salaried-employed jobs, at a particular point in time. The probability of individual i being observed in self-employment ($SE_i = 1$), relative to being in salaried-employment ($SE_i = 0$) can be written as:

$$\Pr[SE_i = 1] = \Pr[SE_i^* > 0] = \Phi(\boldsymbol{x_i}\boldsymbol{\beta} + u_i) \quad [1]$$

where $\boldsymbol{x_i}$ is a vector of strictly exogenous observed characteristics of the individuals, such as age, gender and education level etc.; $\boldsymbol{\beta}$ is the vector of coefficients associated

with $\boldsymbol{x_i}$ to be estimated; u_i is the error term; and, Φ is the non-linear probit function.

The static cross-sectional model in Equation 1, however, neglects the possible impacts of unobserved heterogeneity and state-dependence. In order to control for these impacts it is necessary to instead utilise a dynamic random-effects panel probit framework. In this framework the probability of individual i being self-employed ($SE_{i,t} = 1$), at time t, is assumed to be determined by the individual's previous self-employment status, as well as other observed and unobserved individual characteristics:

$$\Pr[SE_{i,t} = 1] = \Pr[SE^*_{i,t} > 0] = \Phi(\gamma SE_{i,t-1} + \boldsymbol{x_{i,t}\beta} + \varepsilon_i + \nu_{i,t}) \quad [2]$$

The dynamic model in Equation 2 differs from the static model in Equation 1 in two important respects. First, a one-period lag of the observed status in self-employment at year $t-1$ is included as an explanatory variable, $SE_{i,t-1}$. The estimated coefficient of the lagged dependent variable, γ, measures the extent of the effect of state-dependence of self-employment (i.e. the propensity of individuals' participation in self-employment that is determined by their previous experience in self-employment). Second, the dynamic model controls for the unobserved individual heterogeneity component (ε_i) of the error term ($u_{i,t} = \varepsilon_i + \nu_{i,t}$), which includes any individual-specific characteristics that are unobserved in the data but persist over time (e.g. inherent cognitive or non-cognitive abilities).[15] The presence of unobserved heterogeneity is problematic, however, because of its possible correlation with the observed individual heterogeneity in the current time period as well as all previous time periods, which, in turn, has the potential to bias the coefficient estimates on the explanatory variables ($\boldsymbol{\beta}$), particularly the lagged dependent variable (γ).

The inclusion of unobserved individual heterogeneity in a dynamic framework may generate a spurious correlation between individuals' past experience in self-employment on their current propensity to participate in self-employment (Heckman, 1981a). In the model, the assumption of independence between ε_i and the lagged dependent variable ($SE_{i,t-1}$) no longer holds, and the effect of ε_i cannot be eliminated through the simple application of a fixed-effects or random-effects estimator. It is possible, therefore, for the relationship between unobserved

[15] Assume that $\nu_{i,t} \sim N(0, \sigma_\nu^2)$, and is independent of the observed characteristics.

individual heterogeneity and state-dependence in a dynamic framework to be endogenous, which may bias the estimated effect of genuine state-dependence on the persistence of self-employment. This endogenous relationship is commonly referred to as the 'initial conditions problem' (Heckman, 1981b).

In practice, the initial conditions problem arises typically when using longitudinal data, such as HILDA, because the initial self-employment status ($SE_{i,0}$) observed in the data is already the product of a long established but unobserved sequence of behaviours and decisions, determined by unobserved individual heterogeneity (ε_i) and the histories of individuals' characteristics ($\boldsymbol{x_{i,t}}$) and random luck ($\nu_{i,t}$). To deal with the initial conditions problem, Heckman (1981b) suggested approximating the unmeasured history of outcomes conditional on unobserved individual heterogeneity by separately specifying a reduced-form model of the initial self-employment status, using 'pre-sample' information as explanatory variables (e.g. family background or labour market history), and then estimating the reduced-form model jointly with the dynamic model.[16] The Heckman method, however, is rarely implemented in applied research due to its econometric and computational complexity.

This study instead adopts the Wooldridge (2005) method, which suggests approximating the distribution of unobserved individual heterogeneity (ε_i) conditional on the initial self-employment status ($SE_{i,0}$) and the other exogenous observed characteristics. To approximate the initial conditions history, the Wooldridge method suggests the individual-specific means of the time-varying exogenous observed characteristics, $\bar{\boldsymbol{x}}_{\boldsymbol{i}}$ (also referred to as Mundlak (1978) corrections), and the initial value of the individuals' status in self-employment, $SE_{i,0}$, as explanatory variables, denoted as:

$$\varepsilon_i = \mu_0 + \bar{\boldsymbol{x}}_{\boldsymbol{i}}\boldsymbol{\delta} + \gamma_0 SE_{i,0} + \eta_i \qquad [3]$$

where $\eta_i \sim N(0, \sigma_\eta^2)$ and is independent of x and ν for all i and t. The observed explanatory variables (i.e. $SE_{i,0}$ and $\bar{\boldsymbol{x}}_{\boldsymbol{i}}$) are now allowed to correlate with unobserved individual heterogeneity (ε_i), while remaining uncorrelated with the individual-specific error term (η_i). Substituting Equation 3 into the dynamic model

[16] Similar to the Heckman (1981b) method, Orme (2001) suggests a less complex two-step procedure. Orme's (2001) method is used by Henley (2004) to estimate a dynamic probability model of self-employment for the U.K.

in Equation 2, the probability of individual i being observed in self-employment ($SE_{i,t} = 1$) at time t, relative to being an employee ($SE_{i,t} = 0$) becomes:

$$\Pr[SE_{i,t} = 1] = \Pr[SE^*_{i,t} > 0] = \Phi(\gamma SE_{i,t-1} + \boldsymbol{x_{i,t}}\boldsymbol{\beta} + \gamma_0 SE_{i,0} + \bar{\boldsymbol{x}}_{\boldsymbol{i}}\boldsymbol{\delta} + \eta_i + \nu_{i,t}) \quad [4]$$

As before, the estimated coefficient of the lagged dependent variable, γ, measures the extent of the effect of state-dependence of self-employment. Whereas, the estimated coefficient on the initial observed status in self-employment, γ_0, indicates the importance of the correlation between unobserved heterogeneity and the initial condition. Wooldridge (2005) also recommends that interactions between $SE_{i,0}$ and $\bar{\boldsymbol{x}}_{\boldsymbol{i}}$ are necessary if interactions between $SE_{i,t-1}$ and $\boldsymbol{x_{i,t}}$ are included. The Wooldridge method is then easily implemented using a typical random-effects panel probit estimator under common assumptions, which provides a novel and simple solution to the initial conditions problem in the dynamic model.[17]

6.2 Model specification

From the HILDA data, the sample is restricted to an unbalanced panel of Australian workers in salaried-jobs or self-employment, aged 15 years or over, and not studying full-time. In contrast to the conventional approach in labour economics research, the age range is not restricted to the conventional range of working ages, typically 25-64 years, because the age profile of the transitions into self-employment occur much later in the work-life cycle and self-employed workers continue to work well past the conventional age of retirement (approximately 65 years). The sample for the dynamic model is further restricted by the inclusion of the lagged and initial dependent variables, which excludes the wave at which individuals are first observed in HILDA, as well as any subsequent waves for individuals without two or more consecutive observations. Overall, the sample size for the static cross-sectional (or pooled-panel) model is 86,946 observations, representing 17,502 individuals; whereas, the sample size for the dynamic panel model is 64,960 observations, representing, 11,702 individuals.

The dependent and explanatory variables used in the multivariate analysis are, for the most part, determined by the information provided in the HILDA data. As discussed earlier, the dependent variable indicates the workers in the sample who engaged in an employment arrangement classifiable as self-employment (previously

[17] Assuming $\nu_{i,t} \sim N(0, \sigma_\nu^2)$, and is independent of the observed characteristics.

defined) as a binary (or dummy) variable. Furthermore, the lagged dependent variable and the initial value of the dependent variable, included as explanatory variables in the dynamic model (as prescribed by the Wooldridge method), indicate self-employment status in a similar way.

In addition to the lagged and initial sets of expanded dependent dummy variables, a number of explanatory variables are included. The selection of the explanatory variables in this study are based on the variables used previously in similar studies that examine individuals' participation in self-employment (e.g. Evans & Leighton, 1989; Taylor, 1996; Blanchflower & Oswald, 1998; Hamilton, 2000; Henley, 2004; Uhlendorff, 2006; Taylor, 2011). These variables are intended to capture the effects of age[18], the number of resident dependent children, marital/de facto status, gender (female), education[19], long-term health condition or disability, geographic location of residence[20], home-ownership status[21], ethnic origin[22], unemployment rate in local area of residence[23], rates of self-employment in the industry and occupation of work[24], intergenerational occupational match[25], and labour market experience[26]. Furthermore, to address the initial conditions problem (discussed earlier), individual-specific means (i.e. the Mundlak corrections) for each of the time-varying explanatory variables are also included as explanatory variables in the dynamic

[18] Age and age-squared are included as continuous variables that capture the non-linear relationship between ageing and labour market interaction.
[19] Education is included as a dummy set indicating the highest level of education attainment, broadly classified into university, vocational education, Year 12, and Year 11 or below levels of qualification.
[20] Geographic location is included as a dummy set representing urban, and rural/remote areas.
[21] Home-ownership status is included as a dummy set indicating whether a person owns their property of residence outright, holds a mortgage, holds a rent-buy agreement, pays rent or board, or holds a life-tenure agreement.
[22] Ethnic origin is included as a dummy set indicating whether a person was born an Australian native, or a foreign migrant born in either a main-English speaking country or a non-English speaking country.
[23] The unemployment rate is included as a continuous variable indicating the proportion of unemployment that exists in a respondent's local area, using Small Area Labour Market (SALM) information collated and published on a quarterly basis by the Australian Government's Department of Employment (https://employment.gov.au/small-area-labour-markets-publication). This information is matched to the HILDA data using the respondents' reported Local Government Area (LGA) geographic level (as defined by the Australian Standard Geographical Classification (ASGC), see ABS (1996; 2001; 2006)) of residence and for the quarterly time period closest to the respondents' date of interview.
[24] The self-employment rates are included as two continuous variables indicating the proportion of self-employed workers in a respondents industry and occupation of work. The information on the rates of self-employment is estimated from Australian Bureau of Statistics (ABS) national labour force statistics, and then matched to the HILDA survey data using the respondents' reported industry or occupation of work (at the 1-digit level of the Australian & New Zealand Standard Industrial Classification (ANZSIC) or the Australian & New Zealand Standard Occupation Classification (ANZSCO)).
[25] Intergenerational occupation match is included as one dummy indicating whether a person's current occupation was the same as their father's occupation when the person was aged 14. Paternal occupation information is reported by the respondent and matched to their current occupation of work (at the 2-digit level ANZSCO).
[26] Labour market experience is included as two continuous variables measuring the time a person spent in either unemployment or not-in-the labour force as a proportion of the total number of years since completing full-time education, but prior to entering the HILDA survey.

model.

In contrast to previous studies, the unemployment rate and the self-employment rate (by industry and occupation) explanatory variables are precise aggregate time-series measures, which have been matched to the HILDA data from the national labour statistics. These variables provide exogenous indicators of demand-side factors, and capture the prevailing macroeconomic conditions in which individuals' must make their self-employment participation decision. Also, because information on paternal self-employment status is not available in the HILDA data, the possible effect of intergeneration heritability of self-employment on the individuals' participation decision are instead proxied by the paternal occupation match explanatory variable.

Summary statistics of the dependent and explanatory variables for the samples used in both the static and dynamic model estimations are presented in Table 5. There is very little change in the summary statistics of the sample as it becomes smaller due to the restrictions imposed by the dynamic model estimation.

Table 5: Summary statistics of modelling samples

Dependent & explanatory variables	Static pooled-panel sample						Dynamic panel sample					
	Employee	Self-employed	All				Employee	Self-employed	All			
	Mean	Mean	Mean	S.D.	Min.	Max.	Mean	Mean	Mean	S.D.	Min.	Max.
Lagged dependent (base= Employee [t-1]):												
~ Self-employed *[t-1]*							0.03	0.86	0.18	0.38	0	1
Initial dependent (base= Employee [t=1]):												
~ Self-employed *[t=1]*							0.05	0.70	0.16	0.37	0	1
Demographic characteristics												
Age (in years)	39.0	46.9	40.4	12.9	15	89	40.3	47.4	41.6	12.4	16	89
Age^2	1684.0	2342.3	1797.4	1087.8	225	7921	1777.6	2377.7	1884.8	1068.1	256	7921
No. of resident dependent children	0.83	1.09	0.88	1.14	0	12	0.87	1.10	0.91	1.15	0	12
Marital status (base= single):												
~ Married/defacto	0.66	0.83	0.69	0.46	0	1	0.69	0.83	0.71	0.45	0	1
Long-term health condition (base= none):												
~ Disability/impairment	0.13	0.17	0.14	0.34	0	1	0.13	0.17	0.14	0.34	0	1
Geographic location (base= city/urban):												
~ Regional/remote	0.13	0.29	0.16	0.37	0	1	0.14	0.29	0.16	0.37	0	1
Home Ownership (base= mortgage/rent-buy):												
~ Own	0.22	0.37	0.24	0.43	0	1	0.22	0.37	0.25	0.43	0	1
~ Rent/board	0.29	0.16	0.26	0.44	0	1	0.26	0.15	0.24	0.43	0	1
~ Life-tenure (no-equity)	0.02	0.02	0.02	0.15	0	1	0.02	0.02	0.02	0.15	0	1
Education (base= school non-completer):												
~ University	0.28	0.24	0.27	0.44	0	1	0.29	0.24	0.28	0.45	0	1
~ VET	0.33	0.40	0.34	0.47	0	1	0.34	0.40	0.35	0.48	0	1
~ Yr. 12	0.18	0.12	0.17	0.37	0	1	0.16	0.11	0.15	0.36	0	1
Gender [t=1] (base= male):												
~ Female	0.50	0.34	0.47	0.50	0	1	0.49	0.33	0.46	0.50	0	1
Country of birth [t=1] (base= Australia):												
~ Main English speaking	0.09	0.12	0.10	0.30	0	1	0.09	0.12	0.10	0.30	0	1
~ Non-English speaking	0.10	0.12	0.10	0.31	0	1	0.10	0.11	0.10	0.30	0	1

Table 5 (continued)

Dependent & explanatory variables	Static model sample						Dynamic model sample					
	Employee	Self-employed	All				Employee	Self-employed	All			
	Mean	Mean	Mean	S.D.	Min.	Max.	Mean	Mean	Mean	S.D.	Min.	Max.
Employment/labour market characteristics												
SALM unemployment rate *(%)*	5.5	5.3	5.4	2.7	0.0	27.4	5.3	5.1	5.3	2.6	0.0	27.4
Self-employment rate in industry *(%)*	16.5	29.9	18.8	14.5	0.3	65.0	16.0	29.8	18.5	14.5	0.3	65.0
Self-employment rate in occupation *(%)*	18.4	25.7	19.7	11.0	4.9	49.2	18.7	25.6	20.0	10.9	4.9	48.3
Paternal occupation match (base= no match):												
~ Match	0.05	0.14	0.07	0.25	0	1	0.05	0.14	0.07	0.25	0	1
Share of working-life in unemployment *(%) [t=1]*	4.0	2.0	3.6	10.3	0.0	100.0	3.6	1.9	3.3	9.5	0.0	100.0
Share working-life not-in-the labour market *(%) [t=1]*	12.9	10.4	12.5	20.2	0.0	100.0	12.7	10.0	12.2	19.8	0.0	100.0
Observations	71,970	14,976	86,946				53,357	11,603	64,960			
	82.8%	17.2%	100.0%				82.1%	17.9%	100.0%			

Notes: The samples include individuals aged 15 years or over, and not currently studying full-time.
The individual-specific means of the time-varying explanatory variables are not included in this table
~ indicates a dummy variable set.
[t-1] indicates variables with a one-period lag.
[t=1] indicates time-invariant variables, where the initial value is taken at wave of entry
Source: HILDA Survey, 2001-2011

7 Results

The coefficient estimates for the simple pooled static probit model and the more advanced dynamic random-effects panel probit model are presented in Table 6. Because of the use of non-linear probit estimators in this study, the coefficient estimates of the explanatory variables reported in Table 6 cannot be interpreted directly as marginal effects. To make sense of the coefficient estimates and infer an effect, this study adopts the Average Partial Effects (APE) approach and estimates the difference between counterfactual outcome probabilities, holding the explanatory variable of interest fixed at two different values (e.g. values 0 and 1 for a categorical dummy variable). The difference in the counterfactual predicted probabilities is estimated for each individual in the sample, and the marginal effect is then sample average of the individual differences. The APE estimation for the static cross-sectional (or pooled-panel) probit model can be written as:

$$\Pr\left[SE_{i,t} = 1\right] = N^{-1} \sum_{i=1}^{N} \Phi\left[(\hat{\alpha} + \boldsymbol{x_{i,t}}\widehat{\boldsymbol{\beta}})\right] \quad [5]$$

For the dynamic random-effects panel probit model, however, the estimation of the marginal effects is more complex.[27] To provide comparability with the pooled-panel probit model estimates, the coefficient estimates for a random-effects probit model are re-scaled by an estimate of $\sigma_v/\sigma_u = \sqrt{1-\rho}$ prior to calculating the partial effects.[28] Therefore, the APE estimation for the dynamic random-effects panel probit model becomes:

$$\Pr\left[SE_{i,t} = 1\right] = N^{-1} \sum_{i=1}^{N} \Phi\left[(\hat{\alpha} + \hat{\gamma} SE_{i,t-1} + \boldsymbol{x_{i,t}}\widehat{\boldsymbol{\beta}} + \hat{\gamma}_0 SE_{i,0} + \overline{\boldsymbol{x}}_{\boldsymbol{i}}\widehat{\boldsymbol{\delta}})/(1+\rho)^{1/2}\right] \quad [6]$$

In short, rather than compare the difference of a change for a hypothetical person (e.g. set at sample mean values), the APE method compares the difference of a hypothetical change for a sample of individuals. The corresponding average partial, or marginal, effects (APEs) of the coefficient estimates for each model are presented in Table 7.

[27] The coefficient estimates from a random-effects non-linear probit estimator involve different normalisations in comparison a pooled probit estimator (Arulampalam, 1999). The coefficient estimates from a random-effects probit model are normalised on ($\sigma_v^2 = 1$), while for pooled probit models the coefficient estimates are normalised on ($\sigma_u^2 = 1$). Thus, the coefficient estimates of the explanatory variables for a random-effects probit model are β/σ_v, while for a pooled probit model the estimates are β/σ_u.

[28] where $\rho = corr(u_{i,t}, u_{i,s}) = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + \sigma_v^2)$ for $t, s = 2, \ldots, T; t \neq s$

Table 6: Coefficient estimates of models for self-employment probability

Dependent variable: Pr[Self-employed = 1]	Static pooled-panel probit (1)		Dynamic RE panel probit (2)	
	Coef.	S.E.	Coef.	S.E.
Lagged dependent (base= Employee [t-1]):				
~ Self-employed *[t-1]*			1.763***	(0.05)
Initial dependent (base= Employee [t=1]):				
~ Self-employed *[t=1]*			1.808***	(0.09)
Demographic characteristics				
Age (in years)	0.058***	(0.01)	0.121***	(0.02)
Age^2	-0.000***	(0.00)	-0.001***	(0.00)
No. of resident dependent children	0.058***	(0.01)	0.065**	(0.03)
Marital status (base= single):				
~ Married/defacto	0.217***	(0.03)	0.116*	(0.07)
Long-term health condition (base= none):				
~ Disability/impairment	0.083***	(0.03)	0.115***	(0.04)
Geographic location (base= city/urban):				
~ Regional/remote	0.146***	(0.03)	0.074	(0.09)
Home Ownership (base= mortgage/rent-buy):				
~ Own	0.127***	(0.03)	0.038	(0.05)
~ Rent/board	0.001	(0.03)	-0.027	(0.06)
~ Life-tenure (no-equity)	-0.087	(0.08)	0.095	(0.13)
Education (base= school non-completer):				
~ University	0.083**	(0.04)	-0.434*	(0.23)
~ VET	0.094***	(0.04)	-0.244	(0.17)
~ Yr. 12	0.096**	(0.05)	-0.246	(0.25)
Gender [t=1] (base= male):				
~ Female	-0.120***	(0.03)	-0.045	(0.03)
Country of birth [t=1] (base= Australia):				
~ Main English speaking	0.093**	(0.04)	0.148**	(0.06)
~ Non-English speaking	0.118***	(0.04)	0.211***	(0.06)
Employment/labour market characteristics				
SALM unemployment rate *(%)*	-0.005	(0.00)	0.009	(0.01)
Self-employment rate in industry *(%)*	0.029***	(0.00)	0.022***	(0.00)
Self-employment rate in occupation *(%)*	0.019***	(0.00)	0.013***	(0.00)
Paternal occupation match (base= no match):				
~ Match	0.064	(0.04)	0.121	(0.10)
Share of working-life in unemployment *(%) [t=1]*	-0.004***	(0.00)	-0.002	(0.00)
Share working-life not-in-the labour market *(%) [t=1]*	0.001*	(0.00)	0.000	(0.00)
Constant	-4.017***	(0.14)	-4.718***	(0.23)
$\hat{\rho}$ (rho)			0.477***	(0.02)
Log-likelihood	-31265.75		-10218.32	
No. of observations	86,946		64,960	
No. of respondents			11,702	
Average no. of obs. per respondent			5.6	

Notes: Standard errors are in parenthesis
* denotes coefficients significant at = 10%, ** at = 5% and *** at = 1%
Column (2) includes individual-specific means of the time-varying explanatory variables.
The samples include individuals aged 15 years or over, and not currently studying full-time.
~ indicates a dummy variable set.
[t-1] indicates variables with a one-period lag.
[t=1] indicates time-invariant variables, where the initial value is taken at wave of entry

Source: HILDA Survey, 2001-2011

Table 7: Average partial effects on probability of self-employment ($\Pr[SE = 1]$)

Explanatory variables	Static pooled-panel probit (1)	Dynamic RE panel probit (2)
Lagged dependent (base= Employee [t-1]):		
~ Self-employed *[t-1]*		0.265***
Initial dependent (base= Employee [t=1]):		
~ Self-employed *[t=1]*		0.276***
Demographic characteristics		
Age (in years)	0.005***	0.002***
No. of resident dependent children	0.012***	0.005
Marital status (base= single):		
~ Married/defacto	0.041***	0.009
Long-term health condition (base= none):		
~ Disability/impairment	0.017***	0.009*
Geographic location (base= city/urban):		
~ Regional/remote	0.030***	0.006
Home Ownership (base= mortgage/rent-buy):		
~ Own	0.026***	0.003
~ Rent/board	0.000	-0.002
~ Life-tenure (no-equity)	-0.017	0.008
Education (base= school non-completer):		
~ University	0.017**	-0.034
~ VET	0.019***	-0.019
~ Yr. 12	0.020**	-0.019
Gender [t=1] (base= male):		
~ Female	-0.024***	-0.004
Country of birth [t=1] (base= Australia):		
~ Main English speaking	0.019**	0.012*
~ Non-English speaking	0.024***	0.018***
Employment/labour market characteristics		
SALM unemployment rate *(%)*	-0.001	0.001
Self-employment rate in industry *(%)*	0.006***	0.002***
Self-employment rate in occupation *(%)*	0.004***	0.001***
Paternal occupation match (base= no match):		
~ Match	0.013	0.010
Share of working-life in unemployment *(%) [t=1]*	-0.001***	-0.000
Share working-life not-in-the labour market *(%) [t=1]*	0.000*	0.000

Notes: For dummy variables the effect is that of a discrete change (0 to 1)
* denotes the APEs significant at = 10%, ** at = 5% and *** at = 1%
The APEs for the random-effect probit models are estimated using the correction described by Arulampalam (1999).
Column (2) includes individual-specific means of the time-varying explanatory variables.
The samples include individuals aged 15 years or over, and not currently studying full-time.
~ indicates a dummy variable set.
[t-1] indicates variables with a one-period lag.
[t=1] indicates time-invariant variables, where the initial value is taken at wave of entry

Source: HILDA Survey, 2001-2011

7.1 Sensitivity of the results to the econometrics

Changes in the econometric methodology make a significant difference to the estimates obtained, and tell a different story about the contribution of the influences on self-employment choice. Comparing the estimates from the static pooled-panel probit model (Column (1)) to the estimates from dynamic panel probit model (Column (2)), there are several striking differences when state-dependence and unobserved individual heterogeneity are adequately accounted for.

First, the importance of the observed individual characteristics in determining self-employment is greatly diminished both statistically and economically. As shown in Table 6, the strength of the statistical significance of the estimated coefficients on most of the observed characteristics in Column (1) either weaken or disappear completely in Column (2). Many of the statistically significant characteristics in the static pooled-panel model cease to be strongly significant — such as, marital status, geographic location, home ownership, education level, and gender —; while only a handful of the characteristics remain statistically significant in both models — such as, age and age-squared, number of resident children, disability or impairment status, country of birth status, and the rates of self-employment by occupation and industry of work. Moreover, in economic terms, the significant observed characteristics in the dynamic model exert little influence on the probability of being self-employment. As shown in Column (2) of Table 7, the estimated marginal effects of the remaining statistically significant characteristics become increasingly economically irrelevant, with most tending toward zero.. The absence of significant observable characteristics in the dynamic model is in stark contrast to the breadth and diversity of determinants found to be significant in the existing, predominately static cross-sectional, research.

Second, as shown in Column (2) of Table 6, the unobserved effect (denoted as rho ($\hat{\rho}$)) is strongly significant and accounts for approximately 48% of the unexplained variance of the composite error. Controlling for unobserved heterogeneity in the dynamic panel model highlights just how much remains unexplained by the observed characteristics in the static pooled-panel model. This indicates that the choice of self-employment is determined, in part, by inherent individual characteristics that persist over time — such as, cognitive abilities and non-cognitive traits — but these are difficult to capture in the data and are possibly unmeasurable.

Finally, the inclusion of the lagged-dependent and initial-dependent variables (i.e. self-employment status in years $t-1$ and $t=1$) as an explanatory variables in the dynamic model, Column (2) of Table 6, confirm the findings of the previous research on the state-dependence of self-employment, as well as emphasise the need to adequately account for the initial conditions problem. The coefficient estimates on the lagged-dependent ($\hat{\gamma}$) and initial-dependent ($\hat{\gamma}_0$) variables are both positive and highly statistically significant, and have considerably larger magnitudes than any of

the other observed individual characteristics. The statistical significance of estimated coefficient on the initial-dependent variable, as well as the comparability of its magnitude to the lagged-dependent variable, is also indicative of the correlation that exists between the unobserved individual heterogeneity and the initial condition, which is accounted for by the Wooldridge method. The assumption that the initial value of self-employment status observed in the data is determined exogenously does not hold, and ignoring the initial conditions problem likely distorts the estimated coefficients, particularly the effect of state-dependence, at the expense of the unobserved effect.[29] It is important that the initial conditions problem is addressed in dynamic modelling and the Wooldridge method appears to be an effective treatment.[30]

7.2 Importance of state-dependence

The extent to which the probability of workers' participation in self-employment is attributable to genuine or structural state-dependence is substantial. To interpret the effect of the coefficient estimate on the lagged-dependent variable in the dynamic model, Table 8 presents the predicted counterfactual or transition probabilities, as well as the APE for the state-dependent effect. Relative to being an employee, experience in self-employment *in itself* increases the probability of being self-employed in the following year by 27 percentage points (= 0.37 – 0.10). The size of the marginal effect of state-dependence is also considerably larger than marginal effects of any of the remaining significant observed characteristics – presented in Table 7. Furthermore, the size and significance of the influence of state-dependence corresponds with the findings in the international research discussed earlier. The result also confirms the earlier descriptive findings that indicated self-employment as a permanent state rather than a cyclical one; where workers, once engaged in self-employment, remain for a significant proportion of their working lives.

[29] In addition, a Hausman test comparing the difference in results between an exogenously assumed dynamic model (estimates not reported) and the endogenously assumed dynamic model rejected the null-hypothesis and revealed the results to be systematically different from one another (χ^2 = 1460.43, 1 d.f., p-value<0.000).

[30] An additional test of the stability of the Wooldridge method at approximating the distribution of the unobserved heterogeneity at the initial state was undertaken by re-estimating the dynamic model on a restricted-sample (using only the first 10 waves of the HILDA data rather than first 11 waves), and using a Hausman test to compare the results to estimates from the unrestricted-sample. The Hausman test rejected the null-hypothesis. The coefficient estimates for the dynamic model using the unrestricted-sample were found to be systematically different from the estimates for the restricted-sample model (χ^2 = 86.48, 1 d.f., p-value<0.000) at a statistically significant level. However, the difference is particularly small and almost negligible.

Table 8: Estimated transition probabilities & the APE of state-dependence for self-employment & salaried-employment at t, conditional on status at $t-1$

	Self-employed (t)	Employee (t)
Self-employed ($t-1$)	0.367***	0.633***
	(0.02)	(0.02)
Employee ($t-1$)	0.102***	0.898***
	(0.00)	(0.06)
(APE) State-dependence	0.265***	0.265***

Notes: Standard errors are in parenthesis
* denotes the predicted probabilities and APEs significant at = 10%, ** at = 5% and *** at = 1%
Source: HILDA Survey, 2001-2011

The transition probabilities also highlight the importance of state-dependence in determining the self-employment choice in comparison to the obverse choice, the probability of salaried-employment. Independent of the state-dependent effect, the influence of the observed and unobserved characteristics shifts the majority of workers in the sample into salaried-employment rather than self-employment. As shown in Table 8, the combined influence of the observed and unobserved characteristics, independent of employment status in the previous year, only explains 10% of the probability of self-employment in comparison to 63% of the probability of salaried-employment. Therefore, relative to the influence of the observed and unobserved characteristics, state-dependence (27 percentage points) accounts for a considerably greater share of the probability of self-employment ($73\% = 0.27/(0.27 + 0.1)$) than it does for the obverse outcome, the probability of salaried-employment ($30\% = 0.27/(0.27 + 0.63)$). Relative to the salaried-employment opportunities available to self-employed workers, the choice of self-employment is less visible or feasible for employees.

7.3 Separating the effects on transition & survival

Despite the size and the significance of state-dependence on self-employment, it is difficult to discern the economic relevance of its influence. For the reasons discussed earlier, it cannot be said why self-employed workers become locked-in to self-employment (or why employees are locked-out) from the state-dependent effect itself. This is a problem that is only further exacerbated by the absence in the dynamic model of salient differences in worker quality — such as, labour market inexperience or low-levels of education — which could have been used to infer the possible reasons for state-dependence. However, a shortcoming of the estimated effect of the observed characteristics in the dynamic model is that they confound the

probability of self-employed workers *surviving* in self-employment (rather than exiting to salaried-employment) with the probability of employees *transitioning* into self-employment (rather than remaining in salaried-employment). If the transition and survival processes are determined by opposing effects, it is possible that the absence of significant observed characteristics in the dynamic model arises because the net effect of the determinants of transition and survival negate one another.

To disentangle the determinants of transition from those of survival in self-employment, this study extends the dynamic model by interacting the lagged-dependent variable with the observed characteristic variables. Table 9 presents the results of the coefficient estimates for a dynamic random-effects binary model using the Wooldridge method, and including interaction effects, to estimate the probability of being self-employed in year t.[31] In contrast to the previous results, the estimates of the dynamic model with interaction terms separately identify the determinants of the probability transition into self-employment for those who were an employee in year $t-1$, shown in Column (1), from the additional determinants of the probability of survival in self-employment for those who were self-employed in year $t-1$, shown in Column (2). Column (3) reports the tests of joint-significance of the coefficient estimates on each of the characteristic variables and their corresponding interaction terms together. As before, the average partial effects are estimated to interpret the coefficient estimates and are presented in Table 10: Columns (1) and (2) report the effects of the observed characteristics on the probability of self-employment in year t for workers who were either salaried-employed (i.e. transitioned) or self-employed (i.e. survived) in year $t-1$, respectively; while, Column (3) reports the effects of the observed characteristics on the probability of self-employment in year t independent of the workers previous experience.

[31] When interaction effects between the lagged-dependent variable and other observable characteristic explanatory variables are included in a dynamic model using the Wooldridge method, corresponding interaction effects between the explanatory variable for the initial value of the dependent variable and the individual-specific means of the observable characteristic explanatory variables must also be included (Wooldridge, 2005).

Table 9: Coefficient estimates for interaction model

Dependent variable: Pr[Self-employed = 1]	Dynamic binary random-effects probit					
	(1)		(2)		(3)	
			Interaction terms: *Self-employed [t-1]* X		Tests of joint-significance	
	Coef.	S.E.	Coef.	S.E.	χ^2 Coef.	
Lagged dependent (base= Employee [t-1]):						
~ Self-employed *[t-1]*	2.461***	(0.55)	-	-	-	-
Initial dependent (base= Employee [t=1]):						
~ Self-employed *[t=1]*	1.187*	(0.70)	-	-	-	-
Demographic characteristics						
Age (in years)	0.106***	(0.02)	0.013	(0.02)	33.69	***
Age^2	-0.001***	(0.00)	-0.000	(0.00)	20.25	***
No. of resident dependent children	0.080***	(0.03)	-0.037	(0.03)	8.02	**
Marital status (base= single):						
~ Married/defacto	0.132**	(0.06)	-0.055	(0.08)	4.28	
Long-term health condition (base= none):						
~ Disability/impairment	0.101**	(0.05)	0.040	(0.08)	7.96	**
Geographic location (base= city/urban):						
~ Regional/remote	0.182***	(0.07)	-0.266***	(0.09)	14.45	***
Home Ownership (base= mortgage/rent-buy):						
~ Own	0.097*	(0.05)	-0.148	(0.10)	3.56	
~ Rent/board	0.026	(0.07)	-0.166*	(0.09)	4.14	
~ Life-tenure (no-equity)	0.027	(0.14)	0.106	(0.19)	0.57	
Education (base= school non-completer):						
~ University	-0.397	(0.25)	-0.170	(0.11)	5.43	*
~ VET	-0.238	(0.16)	-0.171**	(0.08)	9.05	**
~ Yr. 12	-0.236	(0.26)	-0.304**	(0.13)	6.39	**
Gender [t=1] (base= male):						
~ Female	0.033	(0.04)	-0.157**	(0.07)	5.14	*
Country of birth [t=1] (base= Australia):						
~ Main English speaking	0.150***	(0.05)	-0.016	(0.07)	8.42	**
~ Non-English speaking	0.195***	(0.07)	0.034	(0.10)	13.16	*
Employment/labour market characteristics						
SALM unemployment rate *(%)*	0.010	(0.01)	-0.001	(0.01)	1.10	
Self-employment rate in industry *(%)*	0.027***	(0.00)	-0.015***	(0.00)	227.59	***
Self-employment rate in occupation *(%)*	0.011***	(0.00)	0.003	(0.00)	58.75	***
Paternal occupation match (base= no match):						
~ Match	0.105	(0.09)	0.016	(0.10)	1.57	
Share of working-life in unemployment *(%) [t=1]*	-0.001	(0.00)	-0.011***	(0.00)	8.73	**
Share working-life not-in-the labour market *(%) [t=1]*	0.001	(0.00)	-0.002	(0.00)	0.80	
Constant	-5.084***	(0.30)				
$\hat{\rho}$ (rho)	0.46***	(0.00)				
Log-likelihood	-10070.73					
No. of observations	64,960					
No. of respondents	11,702					
Average no. of obs. per respondent	5.6					

Notes: Standard errors are in parenthesis
* denotes coefficients significant at = 10%, ** at = 5% and *** at = 1%
Estimation include individual-specific means of the time-varying explanatory variables, and subsequent interactions with the initial self-employment status variable.
The samples include individuals aged 15 years or over, and not currently studying full-time.
~ indicates a dummy variable set.
[t-1] indicates variables with a one-period lag.
[t=1] indicates time-invariant variables, where the initial value is taken at wave of entry

Source: HILDA Survey, 2001-2011

In comparison to the previous results, the inclusion of interaction terms to the dynamic model reveals a number of additional observed characteristics that have statistically significant effects on the probability of self-employment (as indicated by the chi-square (χ^2) tests of joint-significance, in Column (3) of Table 9). These additional variables include: rural/regional geographic location, level of education, gender, and the share of time spent in unemployment. Of the jointly-significant variables in Table 9, the estimates also reveal some distinct differences in the observed characteristics that determine the probability of transition to self-employment (Column (1)) from those that additionally determine the probability of survival in self-employment (Column (2)).

Focussing on the significant observed characteristics that determine the probability of employees in year $t-1$ transitioning to self-employment in year t, the results appear to indicate that self-employment attracts employees to whom more beneficial opportunities in salaried-employment may have become less available. The types of employees who are more likely to select into self-employment are those who work in thin' labour markets, such as those living in rural/remote areas or those working in an industry or occupation with a high incidence of self-employment, as well as those with plateauing career trajectories, such as older workers. Further evidence that employees are more likely to transition to self-employment because of a narrowing of employment opportunities is also indicated by the types of employees increasingly attracted to self-employment with characteristics that fall outside of what is typically desired by employers, such as those with long-term disabilities or impairments and migrants (i.e. born in a main-English or non-English speaking country).

Moreover, there is no evidence to support the notion of self-employment as a form of 'disguised unemployment' into which unproductive or poor-quality employees are displaced — such as, those with low-skills or a history of weak attachment to the labour market. None of the variables that capture the skill quality of workers — such as, highest level of education attainment or the acquisition of skills through labour market experience (or lack thereof) — have a statistically significant effect on the probability of employees transitioning to self-employment. Rather than indicate self-employment as an occupation of 'last-resort' for unproductive employees, the results instead reflect the hierarchical nature of organisational structures and the

narrowness of employer hiring practices. The results demonstrate that the skills and labour of the employees who transition still have productive value, and imply that self-employment is an avenue for capable employees to apply their skills and earn a living.

There are also distinct differences between the significant observed characteristics of the employees who enter self-employment and the self-employed workers who stay (or don't exit to salaried-employment), with very little overlap between the same determinants. In contrast to the types of employees entering self-employment, the self-employed workers who accumulate or pool in self-employment appear to be male, poorly educated (i.e. school non-completers), live in a city/urban area, work in an industry with a lower incidence of self-employment, and have spent less time unemployed. However, unlike the entry into self-employment, the results do not provide a consistent story about survival in self-employment, and indicate that the reasons for survival in self-employment may be more complex. A possible explanation is that self-employment is a far more competitive and unforgiving environment than many latent self-employed workers may realise prior to entry.

Despite the statistical significance of the observed characteristics and their corresponding interaction terms, for the most part the economic importance of these variables is trivial. As shown in Table 10, conditional on workers being in either a salaried-employed job (Column (1)) or in self-employment (Column (2)) in the previous year, the marginal effects of the observed characteristics on the current self-employment status are negligibly small. For example, living in a rural/remote area (relative to living in a city/urban area) increases the probability of an employee transitioning to self-employment by 1.7 percentage points, while decreasing the probability of a self-employed worker surviving in self-employment by 1.9 percentage points. The largest marginal effect on the observed characteristics is the effect that an increase in the level of education attainment has on the probability of self-employed workers remaining in self-employment. Relative to having never completed Year 12, self-employed workers with Year 12 or post-school level qualification have a decreased probability of remaining in self-employment by 9 to 12.5 percentage points.

Table 10: Average partial effects for the (interaction) dynamic binary random-effects probit model

Explanatory variables	(1) Conditional on: *Employed [t-1]*	(2) Conditional on: *Self-employed [t-1]*	(3) Unconditional
Lagged dependent (base= Employee [t-1]):			
~ Self-employed *[t-1]*	-	-	0.291***
Initial dependent (base= Employee [t=1]):			
~ Self-employed *[t=1]*	0.248***	0.443***	0.261***
Demographic characteristics			
Age (in years)	0.003***	0.007***	0.002***
No. of resident dependent children	0.007**	0.010	0.005**
Marital status (base= single):			
~ Married/defacto	0.011	0.018	0.009
Long-term health condition (base= none):			
~ Disability/impairment	0.009	0.033	0.009**
Geographic location (base= city/urban):			
~ Regional/remote	0.017*	-0.019	0.007
Home Ownership (base= mortgage/rent-buy):			
~ Own	0.009	-0.012	0.003
~ Rent/board	0.002	-0.032	-0.003
~ Life-tenure (no-equity)	0.002	0.031	0.005
Education (base= school non-completer):			
~ University	-0.033	-0.125*	-0.037
~ VET	-0.020	-0.091*	-0.024
~ Yr. 12	-0.019	-0.119	-0.029
Gender [t=1] (base= male):			
~ Female	0.003	-0.029	-0.002
Country of birth [t=1] (base= Australia):			
~ Main English speaking	0.014**	0.031	0.012*
~ Non-English speaking	0.018**	0.054**	0.017***
Employment/labour market characteristics			
SALM unemployment rate *(%)*	0.001	0.002	0.001
Self-employment rate in industry *(%)*	0.002***	0.003***	0.002***
Self-employment rate in occupation *(%)*	0.001***	0.003***	0.001***
Paternal occupation match (base= no match):			
~ Match	0.009	0.028	0.009
Share of working-life in unemployment *(%) [t=1]*	-0.000	-0.003**	-0.000
Share working-life not-in-the labour market *(%) [t=1]*	0.000	-0.000	0.000

Notes: For dummy variables the effect is that of a discrete change (0 to 1)
* denotes the APEs significant at = 10%, ** at = 5% and *** at = 1%
The APEs for the random-effect probit models are estimated using the correction described by Arulampalam (1999).
Estimation include individual-specific means of the time-varying explanatory variables, and subsequent interactions with the initial self-employment status variable.
The samples include individuals aged 15 years or over, and not currently studying full-time.
~ indicates a dummy variable set.
[t-1] indicates variables with a one-period lag.
[t=1] indicates time-invariant variables, where the initial value is taken at wave of entry
Source: HILDA Survey, 2001-2011

Instead of providing a greater insight into the possible causes underlying the state-dependence of self-employment, the inclusion of the interaction terms in the dynamic model have very little impact on the size of the influence of state-dependence and unobserved heterogeneity. As before, controlling for the impact of state-dependence, unobserved heterogeneity and the initial conditions problem

accounts for most of the influence on the self-employment choice. As evident in Column (1) of Table 9, the coefficient estimates on the lagged-dependent variable ($\hat{\gamma}$), the initial-dependent variable ($\hat{\gamma}_0$) and the unobserved effect ($\hat{\rho}$) all remain strongly significant both statistically and economically. Rather than account for more of the observed heterogeneity, the inclusion of the interaction terms instead slightly increase the size of the influence of state-dependence, while adding very little to the predictive power of the model. In comparison to the dynamic model without interaction terms, there is very little overall change in the mean predicted transition probabilities. As shown in Table 11, the estimated state-dependence for the dynamic model with interactions is 29 percentage points, which is slightly larger than the previous estimate of state-dependence from the dynamic model without interactions (27 percentage points – see Table 8).

Table 11: Estimated state-dependence & the mean predicted transition probabilities for self-employment & salaried-employment at t, conditional on status at $t-1$ (for interaction model)

	Self-employed (t)	Employee (t)
Self-employed ($t-1$)	0.395***	0.605***
	(0.02)	(0.04)
Employee ($t-1$)	0.104***	0.896***
	(0.00)	(0.03)
(APE) State-dependence	0.291***	0.291***
	(0.02)	(0.02)

Notes: Standard errors are in parenthesis
* denotes the predicted probabilities and APEs significant at = 10%, ** at = 5% and *** at = 1%
Source: HILDA Survey, 2001-2011

8 Conclusions

This study shows that modelling the probability of self-employment is extremely sensitive to changes in the econometric method of analysis from a static cross-sectional model to a dynamic panel model. In comparison to the estimates from the static cross-sectional model of self-employment, the results from the dynamic panel estimation greatly diminish the importance of observed heterogeneity in determining self-employment, while revealing the importance of controlling for state-dependence and unobserved heterogeneity, as well as the necessity for dealing with endogenous selection into the initial self-employment state. Consistent with findings from the handful of existing dynamic studies on self-employment for the U.K. and Europe, the results of the dynamic panel model show that self-employment in Australia is also a genuinely persistent state. That is, *ceteris paribus*,

the impact of past experience of self-employment itself increases the current probability of self-employment by 27 percentage points, compared with those who were employees. The influence of genuine state-dependence is also considerably more important in determining the probability of self-employment than it is in determining the obverse outcome, the probability of salaried-employment.

The results of this study are in stark contrast to the importance placed on the observed characteristics in much of the, predominately cross-sectional, evidence on self-employment, and casts doubt on the validity of the findings in the existing research. However, having identified the importance of the influence of genuine state-dependence on self-employment, the possible causes underlying the state-dependence remain unresolved. Unlike like other labour market states, the obscure nature of self-employment and its outcomes make it difficult to interpret or assign meaning to the effect of genuine state-dependence. It cannot be determined from the state-dependent effect itself why self-employed workers become locked-in to self-employment (or why employees are locked-out).

An extension of the dynamic model with interaction terms to disentangle the determinants of transition to self-employment from those of survival provides little further insight. Statistically, the observed characteristics that determine the transition of employees into self-employment are found to be quite different to those that determine the survival of self-employed workers in self-employment. However, for the most part, these differences are economically negligible. Overall, the inclusion of interaction terms in the dynamic model adds very little to the predictive power of the model and does very little to help understand the influence of state-dependence. The possible reasons underlying the influence of state-dependence on self-employment and the transition probabilities continue to remain a mystery.

9 References

ABS. (1996) Australian Standard Geographical Classification. *Australian Bureau of Statistics: Canberra* Catalogue No. 1216.0.

ABS. (2001) Australian Standard Geographical Classification. *Australian Bureau of Statistics: Canberra* Catalogue No. 1216.0.

ABS. (2006) Australian Standard Geographical Classification. *Australian Bureau of Statistics: Canberra* Catalogue No. 1216.0.

ABS. (2013) Labour Statistics: Concepts, Sources and Methods. *Australian Bureau of Statistics: Canberra* Catalogue No. 6102.0.55.001.

Andersson P and Wadensjö E. (2007) Do the unemployed become successful entrepreneurs? *International Journal of Manpower* 28: 604-604.

Arulampalam W. (1999) A Note on Estimated Coefficients in Random Effects Probit Models. *Oxford Bulletin of Economics and Statistics* 61: 597-602.

Benz M and Frey BS. (2008a) Being Independent Is a Great Thing: Subjective Evaluations of Self-Employment and Hierarchy. *Economica* 75: 362-383.

Benz M and Frey BS. (2008b) The Value of Doing What You Like: Evidence from the Self-Employed in 23 Countries. *Journal of Economic Behavior and Organization* 68: 445-455.

Blanchflower DG. (2000) Self-Employment in OECD Countries. *Labour Economics* 7: 471-505.

Blanchflower DG. (2004) Self-employment: more may not be better. *Swedish Economic Policy Review* 11: 15-73.

Blanchflower DG, Oswald A and Stutzer A. (2001) Latent Entrepreneurship across Nations. *European Economic Review* 45: 680-691.

Blanchflower DG and Oswald AJ. (1998) What Makes an Entrepreneur? *Journal of Labor Economics* 16: 26-60.

Buddelmeyer H, Lee W-S and Wooden M. (2010) Low-Paid Employment and Unemployment Dynamics in Australia*. *Economic Record* 86: 28-48.

Cai L. (2014) State-Dependence and Stepping-Stone Effects of Low-Pay Employment in Australia. *Economic Record* 90: 486-506.

Caliendo M and Uhlendorff A. (2008) Self-Employment Dynamics, State Dependence and Cross-Mobility Patterns. *IZA Discussion Paper* No. 3900.

Cappellari L and Jenkins SP. (2008) Estimating Low Pay Transition Probabilities Accounting for Endogenous Selection Mechanisms. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 57: 165-186.

Clark K and Drinkwater S. (2000) Pushed Out or Pulled In? Self-Employment among Ethnic Minorities in England and Wales. *Labour Economics* 7: 603-628.

Covick O. (1998) Self-Employment as Disguised Unemployment. *Unemployment in theory and practice. Lange, Thomas, ed., Cheltenham, U.K. and Northampton, Mass.: Elgar; distributed by American International Distribution Corporation, Williston, Vt.*. 100-121.

Dunn T and Holtz-Eakin D. (2000) Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links. *Journal of Labor Economics* 18: 282-305.

Earle JS and Sakova Z. (2000) Business Start-ups or Disguised Unemployment? Evidence on the Character of Self-Employment from Transition Economies. *Labour Economics* 7: 575-601.

Evans DS and Leighton LS. (1989) Some Empirical Aspects Of Entrepreneurship. *The American Economic Review* 79: 519-519.

Farber HS. (1999) Alternative and Part-Time Employment Arrangements as a Response to Job Loss. *Journal of Labor Economics* 17: S142-S169.

Hamilton BH. (2000) Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment. *Journal of Political Economy* 108: 604-631.

Heckman JJ. (1981a) Heterogeneity and state dependence. In: Rosen S (ed) *Studies in Labor Markets.* Chicago: University of Chicago Press, 91-140.

Heckman JJ. (1981b) The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In: Manski CF and McFadden D (eds) *Structural Analysis of Discrete Data with Econometric Applications.* Cambridge, Mass.: MIT Press.

Henley A. (2004) Self-Employment Status: The Role of State Dependence and Initial Circumstances. *Small Business Economics* 22: 67-67.

Henley A. (2007) Entrepreneurial aspiration and transition into self-employment: evidence from British longitudinal data. *Entrepreneurship & Regional Development* 19: 253-280.

Hundley G. (2001) Why and When Are the Self-Employed More Satisfied with Their Work? *Industrial Relations: A Journal of Economy and Society* 40: 293-316.

Hyytinen A and Rouvinen P. (2008) The Labour Market Consequences of Self-Employment Spells: European Evidence. *Labour Economics* 15: 246-271.

Hyytinen A and Ruuskanen O-P. (2007) Time Use of the Self-Employed. *Kyklos* 60: 105-122.

Le AT. (1999) Empirical Studies of Self-Employment. *Journal of Economic Surveys* 13: 381-416.

Mundlak Y. (1978) On the Pooling of Time Series and Cross Section Data. *Econometrica* 46: 69-85.

Orme CD. (2001) Two-step inference in dynamic non-linear panel data models. *Unpublished paper, University of Manchester. http://personalpages. manchester. ac. uk/staff/chris. orme/documents/Research% 20Paper s/initcondlast. pdf*.

Parker SC. (2009) *The economics of entrepreneurship*: Cambridge Books.

Parker SC and Belghitar Y. (2006) What Happens to Nascent Entrepreneurs? An Econometric Analysis of the PSED. *Small Business Economics* 27: 81-81.

Parker SC and Rougier JC. (2007) The Retirement Behaviour of the Self-Employed in Britain. *Applied Economics* 39: 697-713.

Praag CMv and Versloot P. (2007) What is the value of entrepreneurship? A review of recent research. *Small Business Economics* 29: 351-382.

Rees H and Shah A. (1986) An Empirical Analysis of Self-employment in the U.K. *Journal of Applied Econometrics* 1: 95-108.

Reynolds P, Carter N, Gartner W, et al. (2004) The Prevalence of Nascent Entrepreneurs in the United States: Evidence from the Panel Study of Entrepreneurial Dynamics. *Small Business Economics* 23: 263-284.

Richardson S. (2013) A Reflection on the Household, Income and Labour Dynamics in Australia Survey. *Australian Economic Review* 46: 216-222.

Rotefoss B and Kolvereid L. (2005) Aspiring, nascent and fledgling entrepreneurs: an investigation of the business start-up process. *Entrepreneurship & Regional Development* 17: 109-127.

Shane S and Venkataraman S. (2000) The promise of entrepreneurship as a field of research. *Academy of Management. The Academy of Management Review* 25: 217-226.

Sluis Jvd, Praag Mv and Vijverberg W. (2008) Education and Entrepreneurship Selection and Performance: a review of the empirical literature. *Journal of Economic Surveys* 22: 795-841.

Sørensen JB and Chang P. (2006) Determinants of Successful Entrepreneurship: A Review of the Recent Literature. *SSRN eLibrary*.

Stam E, Audretsch D and Meijaard J. (2008) Renascent entrepreneurship. *Journal of Evolutionary Economics* 18: 493-507.

Stephan U and Roesler U. (2010) Health of entrepreneurs versus employees in a national representative sample. *Journal of Occupational and Organizational Psychology* 83: 717-738.

Stewart MB. (2007) The interrelated dynamics of unemployment and low-wage employment. *Journal of Applied Econometrics* 22: 511-531.

Stewart MB and Swaffield JK. (1999) Low Pay Dynamics and Transition Probabilities. *Economica* 66: 23-42.

Taris TW, Geurts SAE, Schaufeli WB, et al. (2008) All day and all of the night: The relative contribution of two dimensions of workaholism to well-being in self-employed workers. *Work & Stress* 22: 153-165.

Taylor MP. (1996) Earnings, Independence or Unemployment: Why Become Self-Employed? *Oxford Bulletin of Economics and Statistics* 58: 253-266.

Taylor MP. (2011) Self-employment flows and persistence: a European comparative analysis. *ISER Working Paper Series*.

Uhlendorff A. (2006) From No Pay to Low Pay and Back Again? A Multi-State Model of Low Pay Dynamics. *IZA Discussion Papers.* Institute for the Study of Labor (IZA).

von Greiff J. (2009) Displacement and Self-Employment Entry. *Labour Economics* 16: 556-565.

Watson N and Wooden M. (2002) The HILDA Survey: What's in it for Economists? *Australian Journal of Labour Economics* 5: 397.

Watson N and Wooden M. (2013) Adding a Top-Up Sample to the Household, Income and Labour Dynamics in Australia Survey. *Australian Economic Review* 46: 489-498.

Watson N and Wooden MP. (2012) The HILDA Survey: a case study in the design and development of a successful Household Panel Survey. *2012* 3: 13.

Wooden M, Freidin S and Watson N. (2002) The Household, Income and Labour Dynamics in Australia (HILDA)Survey: Wave 1. *Australian Economic Review* 35: 339-348.

Wooden M and Watson N. (2007) The HILDA Survey and its Contribution to Economic and Social Research (So Far)*. *Economic Record* 83: 208-231.

Wooldridge JM. (2005) Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics* 20: 39-54.