LSE Research Online

Christopher Dawson, David de Meza, Andrew Henley and G. Reza Arabsheibani

Entrepreneurship: cause and consequence of financial optimism

Article (Accepted version)
(Refereed)

This version available at: http://eprints.lse.ac.uk/65276/
Available in LSE Research Online: February 2016

http://eprints.lse.ac.uk

Entrepreneurship: Cause and Consequence of Financial Optimism

Christopher Dawson*, David de Meza†, Andrew Henley‡, G. Reza Arabsheibani¶

Abstract

Extant evidence that the self-employed overestimate their returns by a greater margin than employees is consistent with two mutually inclusive possibilities. Self-employment may foster optimism or intrinsic optimists may be drawn to self-employment. Previous research is generally unable to disentangle these effects because of reliance on cross-sectional data. Using longitudinal data, this paper finds that employees who will be self-employed in the future overestimate their short-term financial wellbeing by more than those who never become self-employed. Optimism is higher still when self-employed. These results suggest that the greater optimism of the self-employed reflects both psychological disposition and environmental factors. By providing greater scope for optimism, self-employment entices the intrinsically optimistic.

Keywords: Financial optimism, expectations, self-employment

JEL Classification: D84, M13

* Bristol Business School, University of West of England, † Department of Management, London School of Economics and Political Science, ‡ School of Management and Business, Aberystwyth University and IZA Bonn, ¶ School of Business and Economics, Swansea University and IZA Bonn

Corresponding author: d.de-meza@lse.ac.uk

"The presumptuous hope of success seems…..to entice so many adventurers into those hazardous trades, that their competition reduces their profit below what is sufficient to compensate the risk" Adam Smith, Wealth of Nations (1796), Book 1, Chapter 10.

1. Introduction

Most governments encourage entrepreneurship on the assumption that it is unequivocally a good thing, promoting market competition, innovation and economic growth. For individuals, the attractions of self-employment include personal autonomy, the opportunity to take risk for financial reward, life-style flexibility and tax avoidance (e.g. Scheinberg and McMillan, 1988; Dennis, 1996; Amit et al., 2001; Douglas and Shepherd, 2002; Cassar, 2007). If, as Hamilton (2000) finds, median self-employment income is lower than that in paid employment, or as Moskowitz and Vissing-Jorgenson (2004) report, starting a business involves higher risk but lower expected returns than for stock market investment, these may be prices worth paying. As long as nascent entrepreneurs appreciate the trade-offs, there is no particular cause for concern. However, accumulating evidence suggests that the self-employed overestimate their prospects by more than employees do. This does not necessarily mean that entry into self-employment is excessive. Those selecting self-employment may have an optimistic disposition, in which case they will also tend to overestimate their prospects in paid employment. If returns are exaggerated to the same extent in both activities, the choice of employment mode will be the same as if expectations are realistic. It is only if individuals are more optimistic about self-employment than paid-employment that entry will be excessive. As planning and running a new business are activities that involve a high perception of self-control and offer few barriers to fantasy, they may be fertile conditions for optimism.

Previous research fails to disentangle these effects because it addresses potential associations between optimism and choice of employment mode using cross-sectional data. Consequently, the issue of whether a predisposition towards optimism leads individuals into business formation activity or whether business formation encourages over-optimistic expectations remains unresolved. Does the relative optimism of the self-employed reflect their intrinsic psychology or is it a side effect of planning and running a new business? According to de Meza and Southey (1996), both elements must be present if optimism is to be a cause of self-employment. Were it equally easy to be optimistic concerning prospects in self-employment and paid employment there is no particular reason why intrinsic optimists should be especially attracted by self-employment. Investigation of this issue is only possible with longitudinal data.

To investigate these matters, this paper uses data on a large sample of individuals in the UK tracked annually since 1991. This data source is rich in the sense that it allows sequential observation of financial expectations, financial realizations and transitions into and out of self-employment. However, as explained in the paper, the categorical nature of the data presents various research challenges.

The key finding of the paper is that the future self-employed display above average financial optimism even whilst in paid employment and their optimism becomes greater still on becoming self-employed. Those choosing self-employment in the future are therefore intrinsically more optimistic than those who do not. Moreover, self-employment encourages optimism.

The remainder of the paper is structured as follows. Section 2 describes the background to the questions, identifies some problematic research issues, and proposes strategies to address them. Section 3 describes the longitudinal data source and develops the empirical methodology. This is applied in section 4. Section 5 provides final discussion and conclusions. An appendix presents alternative empirical approaches.

2. Background and development of research issues

There is no completely settled definition of an entrepreneur. Someone who starts a business that employs others is certainly an entrepreneur. Nevertheless, this is not entirely distinct from what, say, a self-employed literary agent does. In this paper, self-employed status is used as an indicator of entrepreneurship. This measure is practical, but may be too inclusive. The accountant with a private practice and an office over a shop may not be a hero of creative destruction. However, as Adam Smith said of the family grocer: "He must have all the knowledge, in short, that is necessary for a great merchant, which nothing hinders him from becoming but the want of sufficient capital." (Wealth of Nations, Book. 1, Ch. 10). The self-employed are typically residual income recipients and willingness to embrace this role suggests they have much in common with narrower definitions of entrepreneurs. Self-employment is the most commonly used measure of individual entrepreneurial status where researchers rely on secondary analysis of existing data (e.g. Amit et al. (1995), Burke et al. (2000), Evans and Leighton (1989), Parker (2009) Taylor (1996), Van Praag and Cramer (2001)). While recognising the limitations, 'self-employed' and 'entrepreneur' are used interchangeably in the discussion.

One attribute of entrepreneurs for which there is accumulating evidence is that they overestimate the financial returns to starting a business (see Parker 2009 for a survey). For example, Cooper, Woo, and Dunkelberg (1988) find that the mean estimate by entrepreneurs of the probability that their own business will survive is well in excess of the mean realized probability. According to Landier and Thesmar (2009), entrepreneurs tend to overestimate employment expansion and sales growth. These studies do not, however, compare the optimism of entrepreneurs with the general population, leaving the question open as to whether higher optimism is really an intrinsic characteristic of those who choose self-employment.[1] Perhaps optimism is equally the province of those in paid employment. After all, optimism has been identified as one of the most widespread of behavioural biases (De Bondt and Thaler (1995, p. 389)).

At first sight, cross-section studies, such as Arabsheibani et al. (2000), Puri and Robinson (2007, 2009), suggest that entrepreneurs do indeed have higher levels of optimism. The self-employed appear to over-estimate financial prospects as well as being optimistic over other domains including lifespan.[2] Fraser and Greene (2006), using British data for the period 1984-99, find that the self-employed have higher income expectations than employees, but the difference diminishes with experience. However, despite using longitudinal data, they

---

[1] Optimism is taken here as forecasts biased in the favorable direction as opposed to overconfidence, interpreted as excessive forecast precision. (It is therefore logically possible to be an overconfident pessimist.) Hvide (2002) terms what we define as optimism, overconfidence1, and what we denote as overconfidence, as overconfidence2. According to Parker (2009), over-optimism refers to over-estimation of the probability of success, whereas over-confidence is under-estimation of the degree of variation in outcomes. Busenitz and Barney (1987) and Forbes (2005) measure overconfidence by examining the proportion of estimates to trivia questions, such as the length of the Nile, that fall outside self-assessed confidence intervals. Business founders are more confident than managers.

[2] It is of course possible that entrepreneurs really do live longer, because, for example, having to receive orders from others might be bad for life expectancy (Kuper and Marmot, 2003).

do not have data on income realizations, and therefore cannot rule out that expectation differences are justified.

Although these studies are of considerable interest, they do not address whether the observations arise because increased optimism is a consequence rather than a cause of entrepreneurship. Almost everyone may think that starting a successful business is more lucrative than it really is.[3] On this view optimism is not a characteristic of the type of person attracted by entrepreneurship but the creation of the noisy, unpredictable environment in which the self-employed typically operate, for, as Kahneman et al. (1982) argue, optimism thrives under conditions of high uncertainty. Amongst reasons is the 'planning fallacy', identified by Kahneman and Tversky (1979). This is the tendency of people engaged in complex projects to regard best case scenarios as the most likely outcome. Explicit planning actually makes the problem worse, encouraging an 'insider' view which places greater weight on internal operational activity than on shocks originating in the external market. By contrast, 'outsiders' place greater weight on typically realized performance, perhaps by paying closer attention to external information about actual realizations of other entrepreneurs. Consistent with this, Cassar (2010) finds that business start-ups which have gone through a formal planning process have the *least* realistic forecasts. Entrepreneurs as insiders in an uncertain environment, where illusions of control flourish and complex planning is required, may be located in an optimism incubator.

---

[3] Astebro (2003) finds inventors have very negative expected returns. Nevertheless, they persist in seeking commercialization even after receiving credible external advice against doing so (Astebro et al., 2007). The latter studies also finds inventors score higher on tests of optimistic attitudes than do the general population. These results are distinctly interesting, though for present purposes subject to qualification. Inventors may be extreme, forecast errors are not measured and it is possible beliefs are adopted post commitment as forecasts.

One way to address the question of whether entrepreneurial optimism is acquired or a prior trait is to compare the financial optimism of people in paid employment who never become self-employed (hereafter '*nevers*') with those currently in paid employment who subsequently become self-employed (hereafter '*futures*'), those in their last period in wage employment (so the forecast is for the first period of self-employment) ('*switchers-in*'), and those in their last year of self-employment ('*switchers-out*') or have been self-employed in the past ('*pasts*'). These groups further contrast with those currently in a spell of self-employment which has lasted for one than one year and which will last for at least one further year ('*selfs*') Since measurements of forecasts and realizations are taken in a common environment, differences should reflect dispositional financial optimism.

There is however an important qualification to this conclusion. If the transition into self-employment is the result of a history of disappointing outcomes in paid employment, *futures* may be recorded as relatively optimistic in the sense that their forecast errors are higher than average. But this is the consequence of rational learning (in the spirit of Jovanovic, 1982) rather than of heterogeneous psychology. Although rational expectations imply that the expected forecast error is zero irrespective of individual characteristics and history, this does not mean that the expected error conditional on future decisions is zero. Suppose, for example, that everyone initially enters paid employment believing they will most likely earn the average of those with the same educational background and other observable characteristics. An individual doing worse faces a signal extraction problem. A poor realization could be the consequence of a bad draw or may be because (unobservable) intrinsic ability is below average for those with similar observables.[4] If learning is rational, a

---

[4] A number of studies have discussed this issue (see, for example, Gervais and Odean, 2001; Camerer and Lovallo, 1999; Cassar and Craig, 2009). Studies suggest, from experimental and other evidence, that individuals are not particularly good at extracting underlying signals of ability from noisy information on forecast performance, and may display self-justification bias.

run of poor realizations should lead to a downward revision of the ability prior and hence of expectations. It may then be worth trying a different way of earning a living, such as self-employment. According to this explanation, the prior optimism of *futures* is associated with worse realizations and forecasts that are no better than those of *nevers* and gradually become worse.[5] If decision making is rational, post-transition into self-employment, the expected forecast error would be zero.[6]

In summary, if those entering self-employment are characterised by greater prior intrinsic optimism, two features should be present. First, those who will be self-employed in the future should over forecast their returns relative to those who will never be self-employed. Second, if this apparent optimism is not (entirely) due to rational learning, the future self-employed should have significantly higher expectations, not just lower realizations. Both of these features readily lend themselves to empirical investigation with appropriate longitudinal data.

3. Data source and descriptive analysis

The data is from the British Household Panel Survey (BHPS). This is a nationally representative general purpose survey funded by the UK Economic and Social Research Council, and similar in structure to the US Panel Study of Income Dynamics and the German Socio-Economic Panel. A stratified random cluster sample of households is drawn from the

---

[5] An alternative rational learning story is that someone who enjoys a good realization run concludes that they are better than average and decides, in the absence of a better paid employment match, to become self-employed as returns are more closely attuned to ability. This though implies that *futures* would be recorded as less optimistic than *nevers*.

[6] If expectations are rational the unconditional forecast error of those with worse past realizations should be zero. In fact the data used in this paper reject this strongly. Furthermore, the unconditional error of those entering self employment should be zero.

population of British household postal addresses in Great Britain.[7] The original sample of 5000 households (approximately 12000 individuals) was recruited in 1991, and follow-on rules are established to track newly forming households involving originally-enumerated household members.[8] The survey instrument is a questionnaire involving a household section, and individual sections, covering a range of topics including economic activity and finances, administered to all adult household members (including new household members at each wave). Repeat interviews take place annually, with 18 annual waves currently available to researchers.

Self-employment in the UK is defined by tax status – that is, registered with the tax authority as an own-account worker or business owner with approval to pay income tax (and social security contributions) through an end of year assessment, rather than through the UK 'pay-as-you-earn' (PAYE) system. The BHPS asks individuals to self-report their employment status, thus identifying self-employment on this basis. *Switchers* and *futures* are identified by those who report that their full-time or main economic status changes. *Switchers-in* and *–out* are those who report a transition into or out of self-employment between the present and the next year. *Futures* are those who report further in the future a change into self-employment.[9] Those who never enter self-employment are identified as *nevers*.

---

[7] The far north of Scotland is excluded because of the prohibitive sampling costs. The original survey excludes Northern Ireland. Booster samples for Wales and Scotland recruited in 1999 and a sample for Northern Ireland recruited in 2001 are excluded from the analysis.

[8] Sample attrition rates in the BHPS are generally low and certainly comparable to those achieved in other similar household panels. As is typical with household panels the highest attrition rate of individuals was between Waves 1 and 2 (12%). Attrition between Waves 2 and 3 was 7% of the original individuals and subsequently averaged 2.4% of the original sample between waves. In common with nearly all previously published research using this data source, attrition is assumed to be a random event.

[9] A small number of transitions into part-time self-employment alongside full-time or part-time paid employment are excluded from the *futures* and *switchers* groups.

From Wave 8 of the survey (1998/9) onwards, all economically active adults are asked about their entrepreneurial aspirations as part of the following question:

'I am going to read out a list of things which you may or may not want to happen to your current employment situation. For each one can you please tell me whether you would like this to happen to you in the next twelve months. Would you like to … start up your own business (a new business)?'

Aspires are identified as those currently in paid employment who in the year of forecast answer in the affirmative; *non-aspires* are those who answer in the negative. This question is specifically about start-ups so accords with narrow definitions of entrepreneurship.

The BHPS allows individual optimism to be investigated by information contained within two questions asked of all individuals in each year. These are:

'Looking ahead, how do you think you yourself will be financially a year from now; better than you are now, worse than you are now, or about the same?' (*forecast*)

and

'Would you say that you yourself are better off, worse off or about the same financially than you were a year ago?' (*realization*)

Matching the first question asked at year *t* with the second question asked at year *t+1* provides forecast and the realization of that forecast.

What subjects understand by being well off financially is not straightforward. A further question asks subjects to attribute the main reason for the change. For those with improved realizations, 67% report that an earnings increase is the main reason, followed by

12% who report a fall in outgoings. There is a close match between *nevers* and *futures* but for *selfs* only 63% name earnings as the main factor. For those experiencing worse realizations, 49% report that the reason is higher outgoings, whereas 29% report lower earnings. Again, *futures* and *nevers* are very similar but 41% of *selfs* report lower earnings and 35% higher outgoings. In judging intrinsic optimism, the source of the change in finances is not obviously relevant. In assessing whether experiencing self-employment affects optimism, it is earnings changes that really matter. Brown and Taylor (2006) compare responses to these questions with real and nominal changes in actual income. The results reassuringly report consistency between an individual's forecasting accuracy and the actual changes in their financial situation, and suggest that individual perceptions square with what happens to actual finances.

To the extent that the realization report is a noisy measure of income change, it will be harder to detect optimism change associated with status transitions. It is not obvious that personal expenses should fluctuate more for the self-employed so changes in optimism when people move status is likely to reflect changes in their income assessment. As on the upside income changes are reported as less important for *selfs* and on the downside more important. This suggests that our subsequent analysis is likely to underestimate how optimism changes when people become self-employed.

Tables 1a and 1b provide descriptive statistics on the BHPS data. Means and standard deviations are reported for available individual-year observations on two sample partitions: firstly *aspires* and *non-aspires*, and secondly *futures, switchers-in, selfs, switchers-out, pasts* and *nevers*. The financial forecasts of *futures* exceed the *nevers* but realizations are rather similar. *Selfs* have higher forecasts than *futures* but also lower realizations. *Aspires* have

higher expectations than *non-aspires* but also better outcomes. As in Das and van Soest (1997), Arabsheibani et al. (2000) Souleles (2004) and Balasuriya, Muradoglu, and Ayton (2010), forecasts and realizations may be cardinalized on three-point scales from which a five-point measure of forecast error can be constructed. A natural approach to determine whether optimism differs between groups is to compare this error. Note that the mean 5-point scale forecast error is in the optimistic direction for all groups but *non-aspires* and *nevers* are the least optimistic by some way, followed by *switchers out, pasts, futures, switchers-in, aspires,* and *selfs*.[10] The rest of the paper investigates whether this optimism ranking can be taken at face value.

TABLE 1 NEAR HERE

The remainder of Table 1 provides descriptive statistics for a set of control covariates which will be used in the econometric analysis: age, gender, marital status and highest education. It is of note that *aspires* and *futures* tend to be slightly younger, and that *non-aspires* and *nevers* are much more likely to be female, reflecting the lower proportion of women amongst the stock of self-employed in the UK. Educational attainment is captured in the analysis through a series of dummy variables indicating the highest level of attainment. These are: university or college degree level at undergraduate or postgraduate level; other non-degree higher education including some historic teaching qualifications and nursing qualifications; A-levels or equivalent (post-compulsory examinations taken at 18 as qualifying exams for college or university entrance), GCSE or O-levels (age 16 schooling

---

[10] *Nevers* are identified as never being self-employed during the sample period. Some may enter self-employment later, in which case the tendency is to under record the extent of the optimism difference with *futures*.

attainment qualifications); and no formal qualifications. *Nevers* are less likely to have higher educational qualifications.

4. Econometric strategy

Denote i's forecast at $t-1$ of their income at t as F_{it} and the realization at t as R_{it}. Initially, suppose both are continuous variables. Defining forecast error as $E_{it} \equiv F_{it} - R_{it}$, the rational expectation, $\bar{F}_{it}$, satisfies $R_{it} = \bar{F}_{it} + e_{it}$ where e_{it} is a random error with mean zero reflecting the various shocks that can intervene between the forecast and realization. It therefore follows that $E_{it} = (F_{it} - \bar{F}_{it}) - e_{it}$. Forecast error is an unbiased but noisy estimate of optimism. The central test therefore appears to be whether the mean forecast error of *futures* significantly differs from *nevers.*

In the present context, there are three problems with this procedure;

i)   Suppose that individuals learn about their productivity over time. Initially, they assume that they resemble their peers and so forecast average returns. As experience flows in, individuals update their estimate of their intrinsic ability. This involves a signal extraction problem. A poor realization may be bad luck or reflect low ability so it is rational to downgrade expectations. Those experiencing a run of poor realizations will therefore be recorded as optimists. After a while, they downgrade their expected returns in paid employment by so much they try self-employment. On average, *futures* are therefore measured as optimistic relative to *nevers,* despite both groups being equally rational. The potential problem is that *futures* are created by subsequent forecast error. In effect, the label

12

is awarded for being an optimist. This rational learning possibility can though be rejected if *futures* have significantly higher realizations.[11]

It is worth making the rational learning issue explicit with a stylized illustration. Let there be equal numbers of two types of risk-neutral people, *A*s and *B*s. If an *A* enters paid employment they earn 100 whereas a *B* earns 50. In self-employment, both types have expected earnings of 60. At the outset people do not know their type. Everyone is risk neutral and has rational expectations. A working life comprises two periods. In the first, everyone chooses paid-employment in which expected earnings are 75, exceeding the 60 in self-employment. As earnings in paid-employment reveal type, in the second period all *B*s switch to self-employment and all *A*s remain in paid-employment. When asked to forecast income, subjects report expected value. Measured by forecast error, *futures* are therefore all optimists and *nevers* are pessimists, despite everyone having rational expectations. Come the second round, everyone is a realist.

Note that the problem identified here does not arise with *aspires* because the classification is made simultaneously with the forecast rather than retrospectively.

ii) As realizations may be subject to large idiosyncratic shocks they are noisy estimates of bias. Detecting between-group differences in optimism may therefore be difficult. There is a better way. Define a new error measure, $\overline{E_{it}} = F_{it} - \overline{R_t}$ where $\overline{R_t}$ is the mean realization of the group to which the individual belongs. It

---

[11] The mirror case is if people do better than average, upgrade their estimate of their ability and then switch to self-employment where they believe reward is more closely related to ability. Rational learning of this sort makes *futures* appear relative pessimists but this is not observed in our data.

follows that averaged over the group, E_{it} equals $\overline{E_{it}}$, but the latter has lower variance if $varR_{it} > 2covR_{it}F_{it}$. This inequality certainly holds if forecasts are random or if the variation in realizations is mostly due to random shocks. In effect, the E_{it} measure assumes the rational expectation is the actual realization and $\overline{E_{it}}$ assumes that the appropriate forecast is the group mean. The difference does not affect the point estimate of bias but its precision. $\overline{E_{it}}$ is normally the appropriate measure if the objective is to detect differences in bias between groups.

iii)   When the forecast error procedure is applied to categorical data of the type in the present data, a potential data truncation bias arises. If a five-point measure of forecast error is constructed, it can then be tested whether the mean of this measure differs between groups. To illustrate why this can give rise to misleading results, suppose that for *futures* the most likely outcome is *better,* so this is their rational forecast. Nevertheless, because outcomes are stochastic, *same* and *worse* are sometimes realized. On average, futures will therefore be recorded as optimists. Similarly, *nevers* may rationally predict worse in which case they will be measured as pessimists on average. It may therefore be falsely concluded that *futures* are significantly more optimistic than *nevers* despite both having rational expectations.

To avoid these three problems, an alternative procedure is followed. An ordered probit is run of expectations at *t* on employment status dummies (with *nevers* as the excluded variable) and other time varying and invariant variables of interest. To take into account that differences in expectations may be rational rather than due to optimism, a performance

control is included. This is created by computing the fitted values from a fixed-effect regression of realization at $t+1$ on time-varying employment status dummies, past realizations age and year dummies.[12] The procedure is a more sophisticated version of using the group mean, as discussed in *ii)* to estimate the rational expectation. The primary interest is in the employment status variables.

A number of specific points should be made concerning this procedure. At the second-stage, the role of the fitted values is as a control. If everyone is equally optimistic, all those with the same fitted value should make the same forecast, therefore the difference in forecast by group measures relative optimism. Although the fitted variables control for rational expectations, because of the categorical data, the only restriction imposed by rationality is this coefficient should not be negative.[13]

Forecast is not included at the first-stage. Its inclusion potentially creates bias. Say that *futures* have on average lower fixed effects. Also, performance is increasing in expectations (as will be true if forecasts have some rationality). It follows by construction that at any given level of predicted performance, the *nevers* must have higher average expectations. At the second stage, the *futures* would therefore be found to be more optimistic, but this is an artefact of the procedure when *futures* are included at the first-stage.[14]

---

[12] The problems with nonlinear fixed effects models (e.g. Greene (2004)) lead us to use a linear formulation albeit that the cardinalization that outcomes involve equal increments is somewhat arbitrary.

[13] Suppose two groups. The best performers have a 40% chance of *better* 30% *same* and 30% *worse*. Their expected performance on a 1,0,-1 scale is 0.1. All these types rationally forecast *better*. The worst performers have a 30% chance of *better* 30% *same* and 40% *worse*. Their expected performance on a 1,0,-1 scale is -0.1 and all of them rationally forecast *worse*. So a change of 0.2 in performance generates a change in forecast of 2, a coefficient of 10. If the two groups had chance of *better* of 40% and 50% performance differs but not forecast, so the coefficient would be zero.

[14] Adding forecasts to the first stage has negligible effect on its explanatory power and the fitted values.

In the case of *futures,* there is a problem in drawing conclusions concerning their relative optimism if this group has lower realizations than *nevers* (as revealed by the average first-stage fixed effects of the groups). Under these circumstances the optimism of the *futures* could be due to rational learning. There are ways round this. If the apparent optimism of *futures* is due to rational learning that their ability is relatively low, inclusion of lagged realizations at the first stage should at least in part control for this. People with same history should draw the same conclusions about their ability and therefore display the same measured optimism even if learning is present. Second, if the optimism of *futures* is due to learning, their expectations should not be significantly higher than *nevers*. So when testing for the relative optimism of *futures,* the procedure is supplemented with a separate test of expectations. If *futures* are more optimistic according to the two-stage test and their expectations are higher, their intrinsic optimism is higher.[15]

The Appendix presents alternative methods of estimating differences in optimism and discusses their merits and drawbacks. All procedures yield similar conclusions.

5. Results

Table 2 reports the first stage of the procedure, the fixed effect regressions used to control for possible performance differences between groups. Column (1) compares *aspires*

---

[15] Separately estimating the expectation and realization equations is more straightforward than using the two-stage procedure. The problem is if it is impossible to reject that *futures* or some other optimistic group perform better. This issue is further discussed in the Appendix.

with *non-aspires* and shows that there is no significant difference in the financial realizations of the two groups. Column (2) compares groups on the basis of current and future employment status. *Switchers-out* have the highest realizations, and the difference is statistically significant. They are followed in order by *pasts* and *selfs*. There is a positive effect of previous realization on current realization, as indicated by the negative coefficients on *same* and *worse*.[16] To measure inter-group performance differences, the fixed effects are retrieved and regressed on the time-invariant group dummies with standard errors bootstrapped. Results for this are reported in Table 3. These results show that *aspires* perform significantly better than *non-aspires* (column 1) and *nevers* perform significantly better than *futures* (column 2).

TABLE 2 NEAR HERE

Optimism estimates are reported in Table 4 where the rational expectation control is the fitted values of the first-stage realization equation.[17] Standard errors are bootstrapped and are clustered in order to take account of multiple observations per individual. The coefficients

---

[16] The regressions in Table 2 include a lagged dependent variable, albeit in a categorical form. In panel data, where the number of time observations is particularly low, there is the possibility of bias in the coefficient estimates. However in the present case the average number of observations per individual is 7. Furthermore the purpose of the model in Table 2 is to provide estimates of realizations rather than inference, and to test differences in averages between groups. There is no reason to believe that any bias would affect particular groups differently.

[17] Alternative specifications such as entering the fitted values in quadratic form to allow for non linear effects, the inclusion of interaction terms hardly affect the final optimism estimates, so for ease of interpretation the simpler form is reported.

on the various time-varying and invariant variables indicate differential optimism.[18] Marginal effects on the probability of forecasting *better* and *worse* are also reported.

TABLE 4 NEAR HERE

Our primary concern is with differences in optimism between the different sample groups. However, before these are discussed, we briefly describe other significant optimism effects, revealed in the other covariate coefficient estimates that are incidental to the main themes of the paper. Firstly optimism declines with age. Secondly men are significantly more optimistic than women. This supports previous research (Puri and Robinson, 2007; Arabsheibani et al. 2000). Again, although statistically significant, the effect is not large. Thirdly being married is associated with lower optimism, supporting previous cross-sectional research. The magnitude of this effect is almost exactly the same in size, but with the opposite sign, as that for males. So unmarried males are most optimistic; married females are least optimistic. Fourthly a higher level of educational attainment is associated with lower optimism. Coefficient estimates vary somewhat across the reported specifications. Generally speaking someone with a university degree or college diploma (HND/HNC) reports lower forecast than someone with no educational qualifications. Lower levels of educational attainment are generally not statistically significant. Finally, optimism is higher following a good realization in the previous period, as indicated by the significant negative coefficient estimates at the foot of the table on the 'same' and 'worse' realization. This is consistent with

---

[18] An alternative procedure runs fixed-effects regressions at both stages, retrieves the second stage fixed effects and then runs them on the group dummies. This yields very similar results. Finally, a pooled first-stage probit can be run at the first-stage to generate rational expectation probabilities to be used as controls in a second-stage expectation probit. Again, results are similar.

people attributing success to their own skill and failure to bad luck (Langer and Roth, 1975). Experiencing success encourages hubris.

The second-stage equations show strong evidence that expectations are at least qualitatively rational in the sense that those with most reason to have above average forecasts according to the first-stage equation, are actually more likely to have higher forecasts.[19]

The main focus of the paper involves comparison of optimism by employment status and aspirations. According to Model 1a of Table 4, *nevers* are significantly less optimistic than all other groups. As *futures* have significantly worse realizations than *nevers,* according to Table 3, their apparent optimism could be due to rational learning. If this were the case, *futures* should not have significantly higher expectations than *nevers* when the rational expectation control is dropped. Model 2a shows this is not the case, so it can be concluded that their optimism is not entirely down to rational learning. [20]

Previous studies have found that the self-employed are more optimistic about their financial prospects than employees. The financial optimism of the self-employed is therefore not (entirely) a consequence of being self-employed but in some measure is a dispositional trait. From Model 1b of Table 4, employees who aspire to start a business are more optimistic than those who do not, again indicating that intrinsic optimism is a factor in self-employment. *People who later become self-employed display more optimism whilst still in paid*

---

[19] Due to the categorical data, rational expectations do not imply that the coefficient on the fitted values is unity. An increase in expected performance may not justify any increase in the most likely outcome or a large change.

[20] For most variables the coefficients in these equations are similar to the expectation without the performance control. This reflects the low coefficient on the performance control and the relatively small performance differences. Expectation differences are a good measure of optimism but this can only be determined by doing the two-stage exercise.

employment than those who never become self-employed. Those who express an aspiration to start a business are also more optimistic.

The next important issue is how optimism changes as people move in and out of self-employment. From Table 4 Model 1a, entrants to self-employment, the self-employed, those leaving self-employment and employees who have been self-employed in the past are all significantly more optimistic than *nevers*. That returners to paid employment are more optimistic is further evidence that intrinsic optimism is a factor in self-employment. *Switchers-in* are more optimistic than *futures* but the difference is not significant, possibly because there are relatively few entrant observations. *Selfs* are more optimistic than *futures* at the 1% level. *There is evidence that the financial optimism of futures is greater when they are self-employed.*

Magnitudes are not small. The probability of a *self* forecasting *better* is some 38% higher than a *never* given both are equally likely to experience *better* and share the same observable characteristics. Before becoming self-employed, the difference is some 10%. Figure 1 shows that the absolute probabilities of forecasting *better* and *worse* of individuals with the mean characteristics of the total sample (including likely performance) but differing in employment status or self-employment aspiration.[21] As noted, *futures* are more optimistic than *nevers,* with optimism further increasing as self-employment is entered. Optimism peaks for those established in self-employment, diminishing when exit occurs, with those who have been self-employed having about the same optimism margin as prior to their entry.

---

[21] The marginal probabilities of Models 1a and 1b are evaluated at the sample means with the status dummies set to zero except for the category of interest.

The most obvious explanation why *futures* and *aspires* are more optimistic concerning their financial prospects than those who remain in paid employment and are happy to do so, follows de Meza and Southey (1996). Optimists are attracted by activities that encourage optimism. For example, individual i's expected return in activity j is $E_{ij} = O_i C_J T_J$ where O_i is an index of I's optimism, C_J the optimism "capacity" of occupation j and T_J is the true return. Defining D_{jk} as the difference in the expected return to two activities, j and k, $\frac{\partial^2 D_{jk}}{\partial O_i \partial C_J} = T_J > 0$. Optimists are characterised by high relative attraction to the activity with provides most fertile conditions for encouraging that optimism. This potentially explains why optimism is higher for *futures* but it implies that their optimism would become greater still when they become self-employed. As noted, there is some evidence that this is the case.

In this analysis optimism matters only because it influences the perception of economic variables. It is possible that optimism is associated with other personality traits that involve a taste for self-employment. Consider 'sensation seeking' defined by Zukerman (1994) as "a trait defined by the need for varied, novel, and complex sensations and experiences and the willingness to take physical and social risks for the sake of such experiences." (p.27) A taste for novelty may be associated with a propensity to give self-employment a go. Nicolaou et al. (2008) find that this characteristic is heritable and entrepreneurs are indeed more likely to have it. In addition, Horvath and Zuckerman (1993) show that sensation seekers underestimate risk. So there is the possibility that the optimism of the self-employed is not a cause of optimism but a side product of the true driver, sensation

seeking. This could be why the financial optimism of the self-employed is not detectably higher than when they were in paid employment.[22]

Whether or not sensation seeking plays a role in explaining the relation between optimism and self-employment, the finding that optimism precedes self-employment has important efficiency implications. Entry due to financial misperception depends on the difference in the optimism applied to paid and self-employment and that is certainly less than cross section estimates suggest.

6. Conclusions

This paper has sought to disentangle the extent to which optimism is a trait predisposing individuals towards self-employment or is a creation of the excitement and intensity that is typically involved in venturing a business. Previous research has established contemporaneous association between over-optimism and self-employment, but, by largely restricting analysis to cross-sectional data, it fails to establish whether over-optimism is a latent characteristic of the future self-employed. Longitudinal data is used to construct derived measures of latent optimism prior to any decision about transition into self-employment.

---

[22] Some suggestive evidence is that sensation seekers tend to be smokers (Zuckerman, Ball, Black, 1990) and, using the BHPS data, we found a strong positive correlation between smoking and financial optimism.

A first key finding is that those who will become self-employed in the future are more optimistic whilst in paid employment than those who will never work for themselves. This finding is robust to a series of investigations about the most appropriate method of measurement. Consistently with this conclusion, employees who aspire to self-employment overestimate their returns by more than those who have no such aspirations. As the desire to start a business is expressed prior to the measurement of optimism, the explanation for this result cannot be rational learning. There is strong evidence that the prior dispositional optimism of the self-employed is above average.

The second key finding is that people become even more optimistic when self-employed. Working for yourself does appear to foster optimism. The cross-sectional difference in the optimism of the self-employed and the paid employed is therefore attributable to both selection and treatment effects. This is a clue to why self-employment attracts optimists. The scope for wishful thinking concerning how well a new business will do plausibly exceeds the opportunity to fantasize about the returns to continuing in paid employment. Natural optimists will therefore exaggerate the difference in returns between self-employment and paid employment, so it is plausible that a track record of optimism in paid employment predicts future self-employment.

It is tempting to conclude from this that entry into self-employment may be excessive. To the extent that optimism affects the assessment of the returns to both paid and self-employment, the existence of optimism bias may not give rise to such distorted entry decisions as may initially appear.[23] As the cross-sectional optimism difference is partly a selection effect, the self-employed also overestimated returns to paid employment. Entry into

---

[23] The fact that 30% of those entering self-employment exit within a year does though indicate error may be involved (Henley 2007).

self-employment is only due to mistaken financial expectations to the extent optimism is lower in paid employment. Cross-section comparisons of the optimism of employees and the self-employed exaggerate the extent of the bias.

References

Amit, R., Muller, E. and Cockburn, I. (1995) "Opportunity costs and entrepreneurial activity" *Journal of Business Venturing* 10(2) 95–106.

Amit, R., MacCrimmon, K., Zietsma, C. and Oesch, J.M. (2001), "Does money matter? Wealth attainment as the motive for initiating growth-orientated technology businesses", *Journal of Business Venturing*, 16(2): 119-143.

Arabsheibani, G., de Meza, D., Maloney, J. and Pearson, B. (2000), "And a vision appeared unto them of a great profit: Evidence of self-deception among the self-employed", *Economics Letters*, 67(1): 35-41.

Åstebro, T. (2003) "The return to independent invention: Evidence of unrealistic optimism, risk seeking or skewness loving?" *Economic Journal* 113 (484):226-239.

Åstebro, T., Jeffrey, S. A., and Adomdza, G. K. (2007) "Inventor perseverance after being told to quit: The role ofcognitive biases." *Journal of Behavioral Decision Making,* 20(3): 253–272.

Balasuriya, J., Muradoglu, Y.G., and Ayton, P. (2010), "Optimism and portfolio choice", working paper, available at SSRN: http://ssrn.com/abstract=1568908 (accessed March 25, 2013)

Brown, S. and Taylor, K. (2006), "Financial expectations, consumption and saving: a microeconomic analysis", *Fiscal Studies*, 27(3): 313–338.

Burke, A. E., FitzRoy, F.R, and Nolan, M.A. (2000) "When less is more: Distinguishing between entrepreneurial choice and performance" *Oxford Bulletin of Economics and Statistics* 62(5): 567–587.

Busenitz, L.W. and Barney, J.B. (1997), "Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making", *Journal of Business Venturing,* 12(1): 9–30.

Camerer, C. and Lovallo, D. (1999), "Overconfidence and excess entry: an experimental approach", *American Economic Review* 89(1): 306-318.

Cassar, G. (2007), "Money, money, money? A longitudinal investigation of entrepreneurial career reasons, growth preferences and achieved growth", *Entrepreneurship and Regional Development*, 19(1): 89-107.

Cassar, G. (2010), "Are individuals entering self-employment overly-optimistic? An empirical test of plans and projections on nascent entrepreneur expectations", *Strategic Management Journal*, 31(8): 822-840.

Cassar, G. and Craig, J. (2009), "An investigation of hindsight bias in nascent venture activity", *Journal of Business Venturing*, 24(2): 149-164.

Cooper, A.C., Woo, C.Y. and Dunkelberg, W.C. (1988), "Entrepreneurs' perceived chances for success", *Journal of Business Venturing*, Elsevier, 3(2): 97-108.

Das, J.W.M., and Van Soest, A. (1997) "Expected and realized income changes: Evidence from the Dutch Socio-Economic Panel", *Journal of Economic Behavior and Organization*, 32(1): 137-154.

De Bondt, W.F.M and Thaler, R.H. (1995), "Financial decision making in markets and firms: A behavioural perspective", in (R. Jarrow *et al..*, eds.) *Handbooks in Operations Research and Management Science*, Vol. 9, Ch.13, 385–410, Amsterdam: Elsevier Science.

de Meza, D. and Southey, C. (1996), "The borrower's curse: Optimism, finance and entrepreneurship", *The Economic Journal*, 106(435): 375-386.

Dennis, J. (1996), "Self-employment: When nothing else is available?", *Journal of Labour Research*, 17(4): 645-661.

Douglas, E.J. and Shepherd, D.A. (2002), "Self-employment as a career choice: attitudes, entrepreneurial intentions, and utility maximization", *Entrepreneurship Theory and Practice*, 26(3): 81-90.

Evans, D.S. and Leighton, L.S. (1989) "Some empirical aspects of entrepreneurship" *American Economic Review* 79(3): 519–535.

Fraser, S. and Greene, F.J. (2006), "The effects of experience on entrepreneurial optimism and uncertainty", *Economica*, 73(290): 169-192.

Forbes, D.P. (2005), "Are some entrepreneurs more overconfident than others?", *Journal of Business Venturing*, 20(5): 623-640.

Gervais, S. and Odean, T. (2001), "Learning to be overconfident", *The Review of Financial Studies*, 89(1): 1-27.

Greene, W. (2004), "The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects", *The Econometrics Journal*, 7(1): 98–119.

Hamilton, B.H. (2000), "Does entrepreneurship pay: An empirical analysis of the returns of self-employment", *The Journal of Political Economy*, 108(3): 604-631.

Henley, A. (2007), **"**Entrepreneurial aspiration and transition into self-employment: evidence from British longitudinal data", *Entrepreneurship and Regional Development* 19(3): 253-280.

Horvath, P. and Zuckerman, M. (1993) "Sensation seeking, risk appraisal, and risky behaviour" *Personality and Individual Differences* 14(1): 41–52.

Hvide, H.K. (2002) "Pragmatic beliefs and overconfidence", *Journal of Economic Behavior and Organization*, 48(1): 15-29.

Jovanovic, B. (1982), "Selection and the evolution of industry", *Econometrica*, 50(3): 649-670.

Kahneman, D., Slovic, P. and Tversky, A. (1982), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press: New York.

Kahneman, D. and Tversky, A. (1979), "Prospect theory: An analysis of decision under risk", Econometrica 47(2): 263-291.

Kuper, H. and Marmot, M. (2003), "Job strain, job demands, decision latitude, and risk of coronary heart disease within the Whitehall II study", Journal of Epidemiology and Community Health, 57(2): 147-153.

Landier, A. and Thesmar, D. (2008), "Financial contracting with optimistic entrepreneurs", *The Review of Financial Studies*, 22(1): 117-150.

Langer, E.J., and Roth, J. (1975), "Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task," *Journal of Personality and Social Psychology*, 32(6): 951-955.

Moskowitz, T. J. and Vissing-Jorgensen, A. (2002), "The returns to entrepreneurial investment: A private equity premium puzzle?", *American Economic Review*, 92(4):745-778.

Nicolaou, N., Shane, S., Cherkas, L. and Spector, T.D. (2008) "The influence of sensation seeking in the heritability of entrepreneurship", *Strategic Entrepreneurship Journal,* 2(1), 7–21.

Parker, S.C. (2009), *The Economics of Entrepreneurship*. Cambridge University Press, Cambridge, UK

Puri, M. and Robinson, D.T. (2007), "Optimism and economic choice", *Journal of Financial Economics*, 86(1): 71-99.

Puri, M. and Robinson, D.T. (2009), "The economic psychology of entrepreneurship and family business", unpublished working paper, Fuqua School of Business, Duke University available at http://ssrn.com/abstract=1422140 (accessed March 25, 2013).

Scheinberg, S. and MacMillan, I.G. (1988), "An 11 country study of motivations to start a business", in *Frontiers of Entrepreneurship Research,* (Eds.) Kirchhoff, B., Long, W., McMillan, W.E., Vesper, K. and Wetzel, W.E., Wellesley, MA: Babson College.

Souleles, N.S. (2004), "Expectations, heterogeneous forecast errors and consumption: micro evidence from the Michigan consumer sentiment surveys" *Journal of Money, Credit and Banking* 36(1): 39-72.

Taylor, M. (1996), "Earnings, independence or unemployment: Why become self-employed?" *Oxford Bulletin of Economics and Statistics* 58(2): 253–266.

Van Praag, C.M. and Cramer, J.S. (2001), "The roots of entrepreneurship and labour demand: Individual ability and low risk aversion" *Economica* 68(269) 45–62.

Zuckerman, M., Ball, S., and Black, J. (1990), "Influences of sensation seeking, gender, risk appraisal, and situational motivation on smoking", *Addictive Behaviors* 15(3): 209–220.

Zuckerman, M. (1994) *Behavioral Expressions and Biosocial Bases of Sensation Seeking*, New York: Cambridge University Press.

Alternative methods of comparing the optimism of the various groups are reviewed here and the corresponding estimates presented. All find that *futures* are significantly more optimistic than *nevers* but evidence that the optimism of futures changes when they enter self-employment is weaker.

A natural procedure is to separately estimate the expectation and realization equations. If the group of interest have significantly higher expectations but lower realizations it is more optimistic. Call this method M1. The problem is that unless it can be ruled out that the realizations of the high expectation group are not better no conclusion can be drawn. This is due to the categorical data. If in one group everyone correctly estimates the probabilities of *worse* is 40% and the others as 30% each, they all forecast *worse.* In another group the probabilities of *better* is 40% with the other outcomes 30% each, so all forecast *better.* The expected outcomes only differ by 0.2 but the expectations by 2. At first sight, the second group is more optimistic, but this is an artefact of the categorical data.

To see whether the combinations that allow conclusions to be drawn apply in this data (augmented to include subjects not observed in paid employment), pooled probit realization and expectation regressions including employment status dummies are run. Table A1 reports the result. A one-tailed test rejects that *futures* have better realizations. As the expectation equation shows that *futures* have higher expectations, it can be concluded that they are more optimistic than *nevers.* For *aspires* and other groups, M1 is not applicable.

The second method, M2, involves the construction of forecast errors. Forecasts and realizations are coded on a three point scale thereby creating a five point forecast error scale.

As discussed in the text, there is a potential truncation error if the group found to be more optimistic has better realizations. From Table A2a, this property does not apply to *futures* relative to *nevers* so the very significant optimism difference between these two groups in Table A2b can be taken at face value.[24] *Aspires* do not have significantly different realizations to *non-aspires* (Table A1). It therefore cannot be ruled out that the relative optimism of the *aspires* in Table A2a is due to truncation bias.

A third method, M3, estimates realization conditional on forecast, controls and group dummies with *nevers* the excluded group. If every group was equally optimistic the distribution of outcomes would be the same given the forecast, so if a dummy is significantly negative, that group is relatively optimistic. M3 estimates realization conditional on forecast whereas the text method, M4, estimates forecast conditional on rational expectation. These procedures are not of equal power. Consider an extreme illustration. Suppose that everyone basis expectations on irrelevant factors. So there is no correlation between forecasts and realizations. Nevertheless *futures* have higher expectations than *nevers,* but there is no significant difference between the groups in the determination of realizations. In this case M4 finds that given expected realization, *futures* have higher expectations; so are more optimistic. In contrast, M3 wrongly concludes there is no difference in optimism between the groups. Although *futures* are more likely to forecast *better,* whichever group such a forecast comes from, there will be no difference in the expected outcome. So M3 is a less direct and reliable estimator of optimism than M4.[25] Nevertheless, Tables A3a and A3b show that

---

[24] An alternative procedure to compare forecast errors by status is first to estimate a fixed-effect model with transition dummies, test for their significance and then test whether the fixed effects of *futures* significantly exceed those of *nevers*. Results are consistent with those of Table A2b, strongly indicating that optimism is both a cause and consequence of self-employment but subject to the same potential truncation bias.

[25] Suppose that the expectation formation process is $E_{it} = O_i f(R_{it}, a_{it})$ where a_{it} are factors unrelated to the rational expectation and O_i is an optimism index. If two groups have the same distribution of R_{it} and a_{it} but differ in their distribution of optimism then M4 measures the difference in mean O. The mean change in R required to preserve E when optimism is higher (what M3 measures) depends on the nature of the f function

futures are significantly more optimistic than *nevers* and *aspires* more optimistic than *non-aspires*.

---

and the joint distribution of R and a so will not necessarily equal the mean difference in O. Consider the following stylised example. Subjects must decide whether they are a G or B. knowing there are equal numbers of both types in the population. A ball is drawn with your letter on it but it's hard to read. An unbiased expert is hired to decipher the letter. An assessment is written specifying the most likely letter that is correct 75% of the time. A realist accepts the report's conclusion. A super optimist reviews the reports and converts Bs to Gs. So for the optimist, 50% of claimed Gs really are. This is not so different to the realist's 50%. If the expert is always right optimism causes the accuracy of the forecast to fall from 100% to 50% but if the expert is no better than random optimism has no effect on the forecast conditional outcome.

Table 1a: Descriptive Statistics

Variable	Description	Aspires Mean	Aspires Std. Dev.	Non-Aspires Mean	Non-Aspires Std. Dev.
Financial forecasts and realizations					
Financial forecast (t):					
Better off	*Reference category*	0.473	0.499	0.342	0.474
Same		0.450	0.498	0.581	0.493
Worse off		0.077	0.267	0.077	0.266
3 point scale (*dependent variable*)	-1 if individual financial forecast 'worse off', 0 if 'same' and 1 if 'better off' at t	0.396	0.627	0.265	0.590
Financial realization (t+1):					
Better off	*Reference category*	0.392	0.488	0.340	0.474
Same		0.366	0.482	0.456	0.498
Worse off		0.242	0.428	0.204	0.403
3 point scale (*dependent variable*)	-1 if individual realized 'worse off', 0 if 'same' and 1 if 'better off' at $t+1$	0.150	0.782	0.136	0.725
Financial realization (t):					
Better off	*Reference category*	0.407	0.491	0.380	0.485
Same		0.365	0.481	0.440	0.496
Worse off		0.229	0.420	0.180	0.384
3 point scale	-1 if individual realised 'worse off', 0 if 'same' and 1 if 'better off' at t	0.178	0.777	0.199	0.721
Forecast error:					
5 point scale (*dependent variable*)	Range from -2 to +2 (Forecast t minus Realization $t+1$)	0.246	0.897	0.129	0.815
Demographics					
Age	Years	35.59	10.42	39.46	12.14
Age squared		1375.2	786.8	1704.7	1001.5
Male		0.628	0.483	0.472	0.499
Marital Status					
Married		0.486	0.500	0.582	0.493
Cohabiting		0.233	0.423	0.163	0.370
Widowed/divorced /separated		0.075	0.263	0.079	0.270
Single, never married	*Reference category*	0.206	0.404	0.175	0.380

Educational Attainment

University degree		0.181	0.385	0.181	0.385
Other higher education		0.092	0.289	0.077	0.267
A-levels		0.223	0.417	0.224	0.417
O-Levels/GCSE's		0.378	0.485	0.364	0.481
No qualifications	*Reference category*	0.126	0.332	0.154	0.361
Number of individual-year observations in each group:		3688		28237	
Number of individuals:		1935		6443	

Definitions: *Aspires*: those currently in paid employment who state a desire to start a business in the next 12 months; *Non-aspires*: those currently in paid employment who do not.

Source: authors tabulations from BHPS 1991-2008.

Table 1b: Descriptive Statistics

Variable	Description	Nevers		Futures		Switchers In		Selfs		Switchers Out		Pasts	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Financial forecasts and realizations													
Financial forecast (t):													
Better off	*Reference category*	0.340	0.474	0.418	0.493	0.428	0.495	0.411	0.492	0.433	0.496	0.351	0.477
Same		0.559	0.496	0.479	0.500	0.466	0.499	0.525	0.499	0.492	0.500	0.575	0.494
Worse off		0.101	0.301	0.103	0.304	0.106	0.308	0.064	0.245	0.075	0.263	0.074	0.262
3 point scale (*dependent variable*)	-1 if individual financial forecast 'worse off', 0 if 'same' and 1 if 'better off' at t	0.239	0.619	0.315	0.649	0.322	0.657	0.347	0.596	0.358	0.617	0.276	0.591
Financial realization (t+1):													
Better off	*Reference category*	0.334	0.472	0.366	0.482	0.374	0.484	0.308	0.462	0.391	0.488	0.320	0.467
Same		0.433	0.495	0.368	0.482	0.349	0.477	0.475	0.499	0.417	0.493	0.466	0.499
Worse off		0.233	0.423	0.266	0.442	0.277	0.448	0.216	0.412	0.192	0.394	0.214	0.410
3 point scale (*dependent variable*)	-1 if individual realized 'worse off', 0 if 'same' and 1 if 'better off' at $t+1$	0.102	0.746	0.099	0.789	0.097	0.802	0.092	0.718	0.199	0.738	0.106	0.723
Financial realization (t):													
Better off	*Reference category*	0.366	0.482	0.414	0.493	0.331	0.471	0.347	0.476	0.321	0.467	0.360	0.480

Same		0.424	0.494	0.351	0.477	0.424	0.495	0.443	0.497	0.417	0.493	0.460	0.498
Worse off		0.210	0.407	0.235	0.424	0.245	0.430	0.210	0.407	0.262	0.440	0.180	0.385
3 point scale	-1 if individual realized 'worse off', 0 if 'same' and 1 if 'better off' at $t+1$	0.157	0.743	0.179	0.785	0.086	0.754	0.137	0.734	0.059	0.762	0.180	0.713
Forecast error: 5 point scale (*dependent variable*)	Range from -2 to +2 (Forecast t minus Realization $t+1$)	0.138	0.837	0.216	0.900	0.225	0.905	0.254	0.834	0.160	0.865	0.171	0.845
Demographics													
Age	Years	38.4	12.1	35.1	10.4	37.7	11.2	43.7	11.8	40.9	11.5	43.2	11.3
Age squared		1619.2	984.0	1340.5	772.4	1546.6	897.0	2048.3	1075.0	1801.7	998.8	1990.8	1035.8
Male		0.467	0.499	0.606	0.489	0.637	0.481	0.671	0.470	0.632	0.483	0.639	0.480
Marital Status													
Married		0.589	0.492	0.541	0.498	0.568	0.496	0.663	0.473	0.640	0.480	0.693	0.461
Cohabiting		0.144	0.351	0.170	0.376	0.203	0.403	0.134	0.341	0.168	0.374	0.137	0.344
Widowed/divorced/separated		0.079	0.270	0.056	0.231	0.067	0.249	0.076	0.264	0.070	0.255	0.084	0.278
Single, never married	*Reference category*	0.188	0.391	0.233	0.423	0.162	0.369	0.127	0.333	0.122	0.328	0.085	0.280
Educational Attainment													
University degree		0.154	0.361	0.198	0.399	0.203	0.403	0.165	0.371	0.199	0.399	0.173	0.378
Other higher education		0.075	0.264	0.091	0.287	0.079	0.270	0.074	0.261	0.098	0.297	0.091	0.287
A-levels		0.205	0.404	0.237	0.425	0.219	0.414	0.251	0.434	0.239	0.427	0.228	0.420
O-Levels/GCSE's		0.373	0.484	0.306	0.461	0.326	0.469	0.305	0.461	0.309	0.463	0.337	0.473

No qualifications	Reference category												
		0.192	0.394	0.168	0.374	0.173	0.378	0.206	0.404	0.155	0.362	0.171	0.377
Number of individual-year observations in each group:		51999		3700		876		2911		614		2102	
Number of individuals:		7697		780		726		809		553		524	

Definitions: *Nevers* – those who never choose self-employment; *Futures* – those currently in paid employment who become self-employed in the future; *Switchers In* – those in their last period of paid employment who will become self-employed in the next year; *Selfs* – those currently in self-employment whose spell has lasted for at least one year and will last for at least one further year; *Switchers Out* – those in their last period of self-employment who will switch out in the next year; *Pasts* – those currently in paid-employment who have been self-employed in the past.

Source: authors tabulations from BHPS 1991-2008

Table 2: Fixed Effect Linear probability regression for financial realizations (first-stage equation)

Variable	(1) Realization s $t+1$ Coefficient	P>z	(2) Realizations $t+1$ Coefficient	P>z
Aspires	-0.011	0.529	-	-
Switchers In	-	-	-0.019	0.608
Selfs	-	-	0.060	*0.024*
Switchers Out	-	-	0.159	*0.000*
Pasts	-	-	0.079	*0.014*
Demographics				
Age	0.003	0.895	0.019	0.184
Financial Realizations time t (reference category: 'better')				
'Same'	0.044	*0.000*	-0.046	*0.000*
'Worse'	0.111	*0.000*	-0.053	*0.000*
Observations	31925		62202	
F test	22.12		15.57	
(p-value)	*0.000*		*0.000*	

Notes: All regressions are clustered by individual and include year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table 3: Linear probability regression of fixed effects from Stage 1 (Table 2)

Variable	(1*) Fixed Effects from (1) Coefficient	P>z	(2*) Fixed Effects from (2) Coefficient	P>z
Aspires	0.0318	*0.004*	-	-
Nevers	-	-	0.0648	*0.000*
Observations	31925		62202	
Chi^2	8.29		18.72	
(p-value)	*0.004*		*0.000*	
Adjusted R^2	0.000		0.002	
Root MSE	0.458		0.511	

Notes: Standard errors are bootstrapped and clustered by individual. *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table 4: Ordered probit regressions for financial forecasts conditionals on predicted realizations (second-stage equation)

	Model (1a): Forecast *t*					
Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Worse)	P>z
Predicted financial realizations at *t+1*	0.636	*0.000*	0.232	*0.000*	-0.093	*0.000*
Futures	0.098	*0.002*	0.036	*0.002*	-0.013	*0.001*
Switchers In	0.156	*0.002*	0.059	*0.003*	-0.020	*0.001*
Self-Employed	0.293	*0.000*	0.112	*0.000*	-0.035	*0.000*
Switchers Out	0.199	*0.000*	0.075	*0.001*	-0.025	*0.000*
Pasts	0.122	*0.001*	0.046	*0.001*	-0.017	*0.000*
Choice threshold parameter 1	-2.234					
Choice threshold parameter 2	-0.401					
Log Likelihood	-52975.3					
chi² (p-value)	*0.000*					
Pseudo R²	0.077					
N	62202					
	Model (2a): Forecast *t*					
Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Worse)	P>z
Futures	0.082	*0.007*	0.031	*0.018*	-0.012	*0.012*
Switchers In	0.146	*0.007*	0.055	*0.008*	-0.020	*0.003*
Self-Employed	0.293	*0.000*	0.112	*0.000*	-0.037	*0.000*
Switchers Out	0.263	*0.000*	0.101	*0.000*	-0.034	*0.000*
Pasts	0.123	*0.000*	0.046	*0.001*	-0.017	*0.000*
Choice threshold parameter 1	-2.425					
Choice threshold parameter 2	-0.630					
Log Likelihood	-53970.7					
chi² (p-value)	*0.000*					
Pseudo R²	0.060					
N	62202					

Table 4 (continued):

Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Worse)	P>z
Model (1b): Forecast *t*						
Predicted financial realizations at *t+1*	0.551	*0.000*	0.203	*0.000*	-0.067	*0.000*
Aspires	0.153	*0.000*	0.058	*0.000*	-0.017	*0.000*
Choice threshold parameter 1	-2.307					
Choice threshold parameter 2	-0.368					
Log Likelihood	-26211.4					
chi² (p-value)	*0.000*					
Pseudo R²	0.075					
N	31925					
Model (2b): Forecast *t*						
Aspires	0.141	*0.000*	0.053	*0.000*	-0.017	*0.000*
Choice threshold parameter 1	-2.545					
Choice threshold parameter 2	-0.648					
Log Likelihood	-26733.6					
chi² (p-value)	*0.000*					
Pseudo R²	0.056					
N	31925					

Notes: Models (1a) and (1b) report standard errors that are bootstrapped and clustered by individual. Models (2a) and (2b) report clustered standard errors. All regressions include age, gender, marital status dummy variables, educational attainment dummy variables, financial realizations at time *t* and a set of year dummy variables (coefficients not reported but available on request). ***Bold italic*** indicates significance level (p-value) below 0.05

Appendix

Table A1a: Ordered probit regression for financial realizations measured at time _t+1 (M1)_

Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Same)	P>z	MFX (Worse)	P>z
Aspires	-0.024	0.280	-0.009	0.278	0.002	0.259	0.007	0.284
Financial Realizations time t (reference category: 'better')								
'Same'	-0.357	_0.000_	-0.129	_0.000_	0.028	_0.000_	0.101	_0.000_
'Worse'	-0.632	_0.000_	-0.207	_0.000_	0.005	_0.025_	0.202	_0.000_
Choice threshold parameter 1	-1.426							
Choice threshold parameter 2	-0.166							
Log Likelihood	-32538.1							
chi² (p-value)	_0.000_							
Pseudo R²	0.033							
N	31925							

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). _Italic_ indicates significance level (p-value) below 0.10, **_bold italic_** below 0.05

Table A1b: Ordered probit regression for financial realizations measured at time *t+1 (M1)*

Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Same)	P>z	MFX (Worse)	P>z
Futures	-0.040	*0.074*	-0.014	*0.071*	0.002	*0.041*	0.012	*0.078*
Switchers In	-0.027	0.567	-0.010	0.565	0.002	0.531	0.008	0.571
Selfs	0.022	0.188	0.008	0.190	-0.001	0.211	-0.007	0.185
Switchers Out	0.123	***0.001***	0.045	***0.002***	-0.011	***0.012***	-0.035	***0.001***
Pasts	0.051	***0.014***	0.019	***0.015***	-0.004	***0.028***	-0.015	***0.012***
Financial Realizations time t (reference category: 'better')								
'Same'	-0.358	***0.000***	-0.127	***0.000***	0.018	***0.000***	0.109	***0.000***
'Worse'	-0.666	***0.000***	-0.214	***0.000***	-0.009	***0.000***	0.223	***0.000***
Choice threshold parameter 1	-1.231							
Choice threshold parameter 2	-0.015							
Log Likelihood	-70459.3							
chi² (p-value)	***0.000***							
Pseudo R²	0.040							
N	68659							

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table A1c: Ordered probit regression for financial forecasts measured at time _t (M1)_

Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Same)	P>z	MFX (Worse)	P>z
Aspires	0.145	*0.000*	0.055	*0.000*	-0.038	*0.000*	-0.017	*0.000*
Financial Realizations time t (reference category: 'better')								
'Same'	-0.399	*0.000*	-0.145	*0.000*	0.092	*0.000*	0.054	*0.000*
'Worse'	-0.197	*0.000*	-0.071	*0.000*	0.043	*0.000*	0.028	*0.000*
Choice threshold parameter 1	-2.831							
Choice threshold parameter 2	-0.935							
Log Likelihood	-26722.7							
chi² (p-value)	*0.000*							
Pseudo R²	0.057							
N	31925							

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table A1d: Ordered probit regression for financial forecasts measured at time *t (M1)*

Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Same)	P>z	MFX (Worse)	P>z
Futures	0.085	*0.005*	0.031	*0.006*	-0.019	*0.008*	-0.012	*0.003*
Switchers In	0.170	*0.000*	0.064	*0.001*	-0.041	*0.001*	-0.023	*0.000*
Selfs	0.236	*0.000*	0.089	*0.000*	-0.058	*0.000*	-0.032	*0.000*
Switchers Out	0.260	*0.000*	0.099	*0.000*	-0.066	*0.000*	-0.033	*0.000*
Pasts	0.162	*0.000*	0.061	*0.000*	-0.039	*0.000*	-0.022	*0.000*
Financial Realizations time t (reference category: 'better')								
'Same'	-0.411	*0.000*	-0.148	*0.000*	0.082	*0.000*	0.066	*0.000*
'Worse'	-0.330	*0.000*	-0.116	*0.000*	0.058	*0.000*	0.057	*0.000*
Choice threshold parameter 1	-2.628							
Choice threshold parameter 2	-0.832							
Log Likelihood	-59554.8							
chi² (p-value)	*0.000*							
Pseudo R²	0.059							
N	68659							

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table A2a: Ordered probit regression for forecast errors *(M2)*

Variable	Coef.	P>z	MFX (2)	P>z	MFX (1)	P>z	MFX (0)	P>z	MFX (-1)	P>z	MFX (-2)	P>z
Aspires	0.109	*0.000*	0.014	*0.000*	0.024	*0.000*	-0.010	*0.000*	-0.024	*0.000*	-0.004	*0.000*
Financial Realizations time t (reference category: 'better')												
'Same'	0.039	*0.007*	0.005	*0.007*	0.009	*0.007*	-0.003	*0.007*	-0.009	*0.007*	-0.001	*0.007*
'Worse'	0.393	*0.000*	0.056	*0.000*	0.084	*0.000*	-0.048	*0.000*	-0.081	*0.000*	-0.011	*0.000*
Choice threshold parameter 1	-2.548											
Choice threshold parameter 2	-1.275											
Choice threshold parameter 3	0.217											
Choice threshold parameter 4	1.193											
Log Likelihood	-38131.5											
chi² (p-value)	*0.000*											
Pseudo R²	0.011											
N	31925											

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table A2b: Ordered probit regression for forecast errors *(M2)*

Variable	Coef.	P>z	MFX (2)	P>z	MFX (1)	P>z	MFX (0)	P>z	MFX (-1)	P>z	MFX (-2)	P>z
Futures	0.086	*0.000*	0.011	*0.000*	0.019	*0.000*	-0.008	*0.001*	-0.019	*0.000*	-0.003	*0.000*
Switchers In	0.129	*0.004*	0.018	*0.008*	0.028	*0.003*	-0.013	*0.020*	-0.028	*0.002*	-0.005	*0.001*
Selfs	0.135	*0.000*	0.018	*0.000*	0.029	*0.000*	-0.013	*0.000*	-0.029	*0.000*	-0.005	*0.000*
Switchers Out	0.064	*0.081*	0.008	*0.096*	0.014	*0.081*	-0.006	0.129	-0.014	*0.075*	-0.003	*0.062*
Pasts	0.062	*0.003*	0.008	*0.005*	0.014	*0.003*	-0.005	*0.009*	-0.014	*0.003*	-0.002	*0.002*
Financial Realizations time t (reference category: 'better')												
'Same'	0.027	*0.007*	0.003	*0.007*	0.006	*0.007*	-0.002	*0.007*	-0.006	*0.007*	-0.001	*0.007*
'Worse'	0.326	*0.000*	0.047	*0.000*	0.070	*0.000*	-0.037	*0.000*	-0.068	*0.000*	-0.011	*0.000*
Choice threshold parameter 1	-2.544											
Choice threshold parameter 2	-1.324											
Choice threshold parameter 3	0.124											
Choice threshold parameter 4	1.105											
Log Likelihood	-83630.6											
chi² (p-value)	*0.000*											
Pseudo R²	0.010											
N	68659											

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table A3a: Ordered probit regression for realizations, conditional on forecasts *(M3)*

Dependent Variable: Realizations $t+1$								
			MFX		MFX		MFX	
Variable	Coef.	P>z	(Better)	P>z	(Same)	P>z	(Worse)	P>z
Aspires	-0.048	*0.031*	-0.017	*0.030*	0.004	*0.019*	0.013	*0.034*
Financial Forecasts time t (reference category: 'better'								
'Same'	-0.340	*0.000*	-0.125	*0.000*	0.034	*0.000*	0.091	*0.000*
'Worse'	-0.901	*0.000*	-0.257	*0.000*	-0.054	*0.000*	0.311	*0.000*
Financial Realizations time t (reference category: 'better')								
'Same'	-0.298	*0.000*	-0.107	*0.000*	0.025	*0.000*	0.083	*0.000*
'Worse'	-0.600	*0.000*	-0.197	*0.000*	0.009	*0.000*	0.188	*0.000*
Choice threshold parameter 1	-1.509							
Choice threshold parameter 2	-0.217							
Log Likelihood	-31873.0							
chi² (p-value)	*0.000*							
Pseudo R²	0.053							
N	31925							

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05

Table A3b: Ordered probit regression for realizations, conditional on forecasts *(M3)*

Dependent Variable: Realizations *t+1*								
Variable	Coef.	P>z	MFX (Better)	P>z	MFX (Same)	P>z	MFX (Worse)	P>z
Futures	-0.057	*0.010*	-0.020	*0.009*	0.003	*0.002*	0.017	*0.011*
Switchers In	-0.058	0.223	-0.020	0.216	0.003	0.123	0.017	0.232
Selfs	-0.023	0.165	-0.008	0.163	0.001	0.141	0.007	0.168
Switchers Out	0.076	*0.049*	0.028	*0.053*	-0.006	*0.096*	-0.022	*0.042*
Pasts	0.021	0.303	0.008	0.305	-0.001	0.328	-0.006	0.299
Financial Forecasts time t (reference category: 'better'								
'Same'	-0.337	*0.000*	-0.121	*0.000*	0.024	*0.000*	0.097	*0.000*
'Worse'	-0.830	*0.000*	-0.239	*0.000*	-0.054	*0.000*	0.293	*0.000*
Financial Realizations time t (reference category: 'better')								
'Same'	-0.294	*0.000*	-0.104	*0.000*	0.016	*0.000*	0.088	*0.000*
'Worse'	-0.611	*0.000*	-0.197	*0.000*	-0.004	*0.023*	0.201	*0.000*
Choice threshold parameter 1	-1.315							
Choice threshold parameter 2	-0.069							
Log Likelihood	-69074.9							
chi² (p-value)	*0.000*							
Pseudo R²	0.059							
N	68659							

Notes: All regressions are clustered by individual and include demographic and education controls, and year dummy variables (coefficients not reported). *Italic* indicates significance level (p-value) below 0.10, ***bold italic*** below 0.05
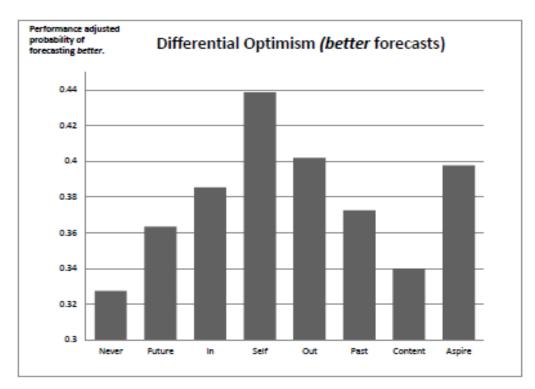
Figure 1: Summary of probability differences between different sample groups