Thoughts on the PSED

Simon C. Parker

University of Durham, Durham, UK,

Background

As Paul Reynolds states in his Overview, new venture creation is the very wellspring of economic dynamism. Entrepreneurial entry is widely promoted, supported, and studied by policy makers, practitioners, and academics.

Individuals who are actively involved in starting a new business are often referred to as "nascent entrepreneurs". The last decade has witnessed the emergence of a new literature that explores the characteristics, development practices and processes, and performance of nascent entrepreneurs. The origins of this literature can be directly traced back to the efforts of Paul Reynolds, Nancy Carter, Bill Gartner and their associates in the early 1990s, which have finally culminated in what we now know today as the PSED.

The positive contribution of PSED

The PSED enables researchers to analyse how opportunities are detected and how new organisations come into being. It overcomes problems of "survival" and "hindsight" biases that arise when data are compiled only for entrepreneurs whose start-ups ultimately survived, and after the start-up itself actually occurred.

As the Overview explains, the PSED has already generated an impressive volume of research, and is sparking a lively research agenda which continues to develop rapidly in a bewildering number of directions. Many detailed descriptive findings from the PSED appear in Gartner *et al*'s (2004) *Handbook of Entrepreneurial Dynamics*, on which the Overview is based; and useful surveys of research that draw heavily or exclusively upon the PSED include Davidsson (2006) and Wagner (2006), as well as articles in the August 2006 Special Issue of *Small Business Economics* devoted to Nascent Entrepreneurship. Instead of commenting on findings generated from the PSED, I will instead consider some of the broader positive contributions made by this dataset, as well as some niggling drawbacks that researchers interested in exploring the dataset further should be aware of before they commence.

From a scientific perspective, perhaps the most valuable contribution of PSED is the generation of a new paradigm for analysing venture start-ups. This explains its extensive and growing use by the research community. The PSED combines a grand ambitious sweep of chronological process relating to new venture creation with a remarkably comprehensive and meticulously detailed compendium of information about numerous issues bound up in the process. This is an irresistible combination, providing versatility, breadth, and depth in a fundamentally relevant and practically useful single entity.

Several aspects of the PSED place it apart from other datasets which are also used to study early stage firm formation. Unlike many of these other datasets, the PSED is dedicated solely to improving our understanding of the phenomenon. It is actually not one dataset but several - and it has been designed on a grand scale. No fewer than 64,000 randomly selected American adults were interviewed in the initial "Screener" stage of the data collection process; this file forms part of the dataset and contains a mine of information which in my opinion remains strangely under-utilised by researchers. Nascent entrepreneurs (and "intrapreneurs") identified at the initial stage were then followed up in a dedicated questionnaire which culminated in the acquisition of over 5000 variables on 1281 individuals. The size of this sample, and further data files that over-sample women and minorities, provide an unparalleled opportunity to perform robust empirical analysis with a meaningful degree of precision and generality. The dataset is further enhanced by returning to respondents in a sequence of follow-up waves, to provide a genuine longitudinal dimension to the PSED as well. All in all, this amounts to an extraordinary data collection exercise - perhaps the most ambitious and impressive of its type ever undertaken in the field.

The practicality of the PSED is another key strength. The dataset provides weights which enable the researcher to draw nationally representative inferences about a wide range of venture creation activities in America. The rich volume of information on many individual-level variables and the sophistication of measurement certainly make PSED the most useful data set with which to perform micro level analysis (Davidsson, 2006), and without doubt preferable to the Global Entrepreneurship Monitor (GEM) in this regard. Last, but not least, the PSED has another compelling feature which is bound to recommend it to researchers. It is well documented, available on line, together with codebooks and questionnaires. And, it is free.

Some limitations

That completes the eulogy. Now for some niggles.

First, the PSED relates to only country: the USA. Unlike GEM, its sister datasets in other countries are not directly comparable, although there are important similarities between PSED and these other datasets. This consideration obviously reduces the appeal of PSED to researchers interested in analysing nascent entrepreneurship outside of the US.

Second, it is important to be clear that despite its emphasis on individual-level factors, "nascent entrepreneurship" in the PSED refers to the venture rather than to the person. In practice, some nascent entrepreneurs are starting a new business for the second, third or more time; and between 5 and 10% of nascent entrepreneurs who were also developing more than one venture at the outset were discarded from the main dataset. There is a related problem here about how the PSED "captures" nascent entrepreneurs, which may include cases who not "genuine" nascent entrepreneurs while excluding individuals who are.

To clarify this point, consider non-genuine nascents. Regular users of the PSED know all too well that it contains several troublesome "nascent entrepreneurs" who claim to have been developing a business for many years, without ever having done anything tangible to progress the situation. More generally, respondents tend to exhibit pronounced heterogeneity with regard to the length of time they claim to have been developing the business. This is clearly an issue that should be treated carefully by researchers intending to use the PSED. One cannot help feeling that some of these "nascent entrepreneurs" are actually no more than dilettantes or dreamers. Whether they should be included at all in any analysis of the PSED is a moot point; but if the researcher does decide to exclude them, they must propose some more or less arbitrary cut-off time beyond which nascent entrepreneurs stop being "doers" and become "dreamers" instead.

The opposite problem, that the PSED does not include every nascent entrepreneur, is suggested by findings from general purpose longitudinal surveys that the majority of transitions into entrepreneurship are not anticipated by individuals even as recently as a year before (Katz, 1990; Henley, 2006). Such individuals would presumably not respond to the PSED interviewers in ways that would identify them as nascent entrepreneurs; yet they start new firms too. This has two practical consequences. First, it means that the PSED likely under-estimates the number of nascent entrepreneurs (the point about dreamers above notwithstanding). Second, if (as seems likely) individuals who develop new businesses very rapidly and are missed by the PSED have different characteristics to individuals who take longer to prepare for entry and are included in the sample, there is a danger that the PSED will give a biased perspective about the nature of nascent entrepreneurship. Of course, this does not mean that researchers ought to use general purpose longitudinal datasets instead of the PSED, as they invariably suffer from a chronic lack of information about the start-up process.

A third limitation is that closer investigation of the panel dimension of the PSED reveals it to be slightly less extensive than first appears. Individuals who did not continue in nascent entrepreneurship beyond the first wave (either because they succeeded in starting a firm or simply quit altogether) were not followed up in detail in subsequent waves. This cuts down the information content of the panel, as we do not know whether individuals ever return to nascent entrepreneurship and if so why and when. The information content is reduced further by high attrition rates, whereby original respondents can no longer be contacted. This does not mean that longitudinal analysis cannot be performed (see, e.g., Parker and Belghitar, 2006), but it does circumscribe what can be done with the data.

Fourth, not all social scientists will find the PSED equally useful. Reflecting the interests and disciplinary backgrounds of the original research consortium members, the PSED addresses itself most satisfactorily to issues in sociology and business/management. Thus there is an abundance of variables measuring intentions, motivations, beliefs etc. In contrast, key constructs in economics, such as risk aversion and intertemporal discounting among others, are absent. To be fair, the PSED still contains plenty of data that do relate to economic issues, as well as zip codes which can be connected to external sources of macroeconomic data. Using these might boost the policy contributions that can be made with the PSED. As the Overview explains, the PSED can inform the policy

debate in a general sense, though there seems to be a limit as to how much *specific* policy analysis can be performed with it. Making connections to policy institutions using linked data sets might help to boost the policy side to research based on the PSED.

Fifth, despite the best efforts of Paul Reynolds and his collaborators, a lot of knowledge about the data set remains tacit. The Overview is valuable in filling some of the gaps, but various subtleties and idiosyncratic coding errors still remain only in the brains of the individuals closely involved with the data collection. Presumably, this information will gently depreciate and become vaguer. Finally, the sheer scale and complexity of the PSED means that it can be daunting to invest time and effort trying to understand it for use in one's own research. Paradoxically, this raises the possibility that the very strengths of the dataset might even work against it!

Finally…

Overall, one can only warmly welcome this Overview which further clarifies the content and context of this wonderful and important dataset, while hopefully advertising its existence to a new audience of potential users. Being aware of the imperfections of the PSED are essential if the dataset is to be used wisely, so the above limitations are intended only in a constructive sense, and definitely not as a criticism of the dataset itself. Without in any way playing down the colossal and sustained efforts of so many collaborators on the PSED project, we have been fortunate indeed to have had such a talented, energetic and intellectually ambitious figure as Paul Reynolds to drive this process forward and create a public good as useful and enlightening as the PSED.

References

Davidsson, P. (2006) Nascent entrepreneurship: Empirical studies and developments, *Foundations & Trends in Entrepreneurship*, 2(1), 1-76.

Gartner, W.B., Shaver, K.G., Carter, N.M. and P.D. Reynolds (2004) *Handbook of Entrepreneurial Dynamics*, California: Sage Publications Inc.

Henley, A. (2006) Entrepreneurial Aspiration and Transition into Self-employment: Evidence from British Longitudinal Data, *Mimeo, Swansea University, Wales*.

Katz, J. (1990) Longitudinal analysis of self-employment follow-through, *Entrepreneurship and Regional Development*, 2, 15-25.

Parker, S.C. and Y. Belghitar (2006) What happens to nascent entrepreneurs? An econometric analysis of the PSED, *Small Business Economics*, 27, 81-101.

Wagner, J. (2006) Nascent entrepreneurs, In *Handbook of Entrepreneurship Research* (ed. S.C. Parker), New York: Springer, 15-37.