



Data replication and extension: A commentary



Per Davidsson ^{a,b,*}

^a Australian Centre for Entrepreneurship Research, Queensland University of Technology, GPO Box 2434, Brisbane 4001 QLD, Australia

^b Jönköping International Business School, Sweden

ARTICLE INFO

Article history:

Received 7 February 2015

Accepted 11 February 2015

Keywords:

Replication

Nascent entrepreneurship

Null findings

Business Planning

ABSTRACT

Honig and Samuelsson (2014) and Delmar (2015) recently had an exchange in this journal related to a replication-and-extension attempt of two papers which originally arrived at different conclusions based on the same data set. This commentary provides further clarification on the issues and links the debate to broader issues scholarly culture and practices in entrepreneurship research.

© 2015 Elsevier Inc. All rights reserved.

Honig and Samuelsson (2014) recently published in this journal an interesting example of re-examining the evidence underlying conclusions regarding the efficacy of business planning for venture creation success. Delmar (2015) has since published a rebuttal. The debate focuses particularly on two previous papers on the efficacy of business planning among nascent ventures. The papers, which I will refer to as “D&S” (Delmar and Shane, 2003) and “H&K” (Honig and Karlsson, 2004), used the same data base but arrived at different main conclusions. I wish to add this commentary because as initiator and leader of the research program from which the underlying data emanates I feel partly responsible for the apparent controversy. I may also be in a unique position to provide further clarification on the debated issues. Further, this case of contradictory findings and replication attempt provides an excellent opportunity to reflect more broadly on the scholarly culture and practices in entrepreneurship research.

I have considerable familiarity with panel studies of nascent entrepreneurship through deep involvement in three major projects and as reviewer of the research stream (Davidsson, 2004b, 2005; Davidsson and Gordon, 2012; Davidsson et al., 2011a, 2011b; Delmar and Davidsson, 2000). The people involved are respected colleagues whose contributions to the nascent entrepreneurship literature I have singled out as particularly valuable (Davidsson and Gordon, 2012: 871–872; Davidsson et al., 2011a, 2011b: xxiv–xxix). With all but one of them I have published research based on the very same data set (Delmar and Davidsson, 2000; Honig et al., 2005; Samuelsson and Davidsson, 2009). I acclaim Shane's scholarship in Davidsson and Wiklund (2009). Further, I devoted a whole chapter in *Researching Entrepreneurship* to the importance of replication (Davidsson, 2004a). As regards the value of business planning for nascent ventures the collective evidence I have seen makes me remain a sceptic.

To put the current debate in perspective it is also important to realize how little was known some 15 years ago about what entrepreneurial processes looked like in a random (-ish) sample, or how they could best be theorized and analyzed. Before scores of researchers sunk their teeth into the PSED and its sister projects, we had very vague notions of the nascent entrepreneurship phenomenon. The dominant, implicit mental model of the phenomenon was arguably that motivated,

* Correspondence address: Australian Centre for Entrepreneurship Research, Queensland University of Technology, GPO Box 2434, Brisbane 4001 QLD, Australia.

E-mail address: per.davidsson@qut.edu.au

single individuals or custom-built teams purposefully entered the process at a distinct point in time with an intention of working full time to implement a well-defined, for-profit venture idea as quickly as possible, following one of a small set of possible types of process, and being at the risk of losing significant amounts of money should they have to abandon the attempt. Early, descriptive results – published or unpublished – proved most such assumptions wrong and revealed a phenomenon of greater complexity, variability and (for the most part) modesty than anyone had thought possible (Aldrich et al., 2006; Carter et al., 2003; Davidsson and Gordon, 2012; Delmar and Davidsson, 2000; Liao et al., 2005; Reynolds, 2007; Reynolds et al., 2004; Ruef et al., 2003; Van Gelderen et al., 2005). We should also remember that we started with limited experience in the particular challenges and opportunities associated with longitudinal, panel data.

In this environment, D&S and H&K were both pioneering studies, providing data, theoretical angles, results, and method solutions that were novel for their time (Davidsson and Gordon, 2012: 871 Davidsson et al., 2011a, 2011b: xxiv–xxix). As a result, they both got published in leading, mainstream management journals – a rare event for entrepreneurship research at the time – and have been widely cited. In order to achieve this, the respective authors had to make a series of methods choices in partly novel terrain. Using the same data set, it is possible that making better and worse choices led one team but not the other to arrive at the correct conclusion on the substantive issue; the efficacy of business planning for nascent ventures.

Do Honig and Samuelsson (2014) settle the issue? I do not think so. Their temporal extension of H&K confirms the original results, but does so without including what we have come to realize is the most important control variable, namely how far progressed the start-up already was when first sampled (Davidsson and Gordon, 2012). Their dissection of D&S reads as if they have identified fundamental and indisputable flaws in D&S' procedure, such that sample bias and reverse causality drive the positive results for business planning. Reading all sides of the arguments, I am not convinced this is the case. The ideal study of the effects of business planning in new venture creation would have what Delmar (2015) calls an "incident cohort design", i.e., capture all cases at a unique and indisputable process inception point and follow them forward over time. As I understand it, what D&S tried to do was to take their study closer to this ideal. They seem to have been the first to realize and try to counteract the sampling and retrospection biases associated with the PSED methodology's inclusion of "long in process" cases that were "still trying" at the time of sampling (while their terminated or operational cohort peers were excluded by the sampling mechanism, i.e., left-censoring). They were also the first to re-organize the data set into monthly spells using the time-stamped gestation activities, rather than relying on the timing of interview waves, which occur at different and arbitrary stages of development across cases. In order to do so, they needed to develop a criterion for process inception. The problem is that because people drift into – and proceed with – start-up processes in all manner of ways there will never exist a distinct and indisputable criterion for when venture creation processes are started (Bhave, 1994; McMullen and Dimov, 2013; Reynolds, 2007; Reynolds and Miller, 1992). Honig and Samuelsson's (2014) analysis indicates that D&S choice of inception criterion may be associated with considerable problems. However, as I understand the set-up of D&S analysis, it is only *change* in planning status associated with *subsequent change* in the respective dependent variables that can affect the analysis results.

By contrast, the H&K analysis approach disregards the fact that "short in process" cases are under sampled, and that the cases are unequally far progressed in the start-up process when they enter the sample. This is understandable as most of us did not realize the importance of this back then (cf. Davidsson and Honig, 2003). However, H&K have one major advantage over D&S, namely that they apply a much stronger criterion for venture creation success than D&S do. The latter rely on indicators that can better be labeled "persistence" and "progress". This is also understandable, because 15 years ago we had limited understanding that this lumps together (a) cases on their way to successful entry; (b) cases that will never be successful and therefore should better be terminated quickly, and (c) cases that tinker about without ever seeming to get to a resolution either way (Davidsson, 2008; 184–185; Reynolds, 2007: 54). H&K's results indicate that planning may be positively associated with persistence but not with venture creation success, which I would say accords with what the collective evidence in nascent entrepreneurship research to date suggests (Davidsson and Gordon, 2012: 858–860). Following up this result after 5 and 10 years, so as to obtain "final outcomes" for a larger share of the sample, is a laudable undertaking. However, due to the methodological limitations of the H&K approach and possible idiosyncrasies of the particular data set I do not think the extended re-analysis settles the substantive issue of the merits of business planning for nascent venture.

This said, there are other important lessons to be learned from their effort and Delmar's rebuttal:

1. It reminds us that *doing good social science is very hard*. In developing their studies, researchers face limitations and are forced to make choices of such nature that close scrutiny may lead to valid questioning of the conclusions. The immense difficulty of the task is reason to respect the honest efforts of our peers, even if we disagree with some of the choices they made. I cannot bring to mind a case where a single study provides rock solid evidence on an issue, no matter how far up the journal hierarchy we look. This means that the current culture in management research, which requires a theoretical contribution of every empirical paper while actively discouraging replication, is fundamentally unsound (Hambrick, 2007; Hubbard and Lindsay, 2013a). We should therefore welcome journals like JBV*i*, which broaden the view of what constitutes valuable scholarly contributions.
2. The difficulty of getting all methods choices right is one reason why *replication* – in a broad sense – *is critically important*. Other reasons are that idiosyncrasies of the original empirical setting may drive the results in a particular direction, and that authors may – consciously or subconsciously – gear their methods choices towards results that match their

hypotheses (Simmons et al., 2011). I do not find reason to think that H&K or D&S are better or worse than others in this regard, but we know that *confirmation bias* is a pervasive human trait and that researchers are not immune (Kassin et al., 2013). As a matter of fact, this bias is institutionally supported by a research culture that is narrowly focused on “statistically significant” results and unappreciative of “null” findings (Hubbard and Lindsay, 2013b; Landis and Rogelberg, 2013; O’Boyle et al., 2014; Simmons et al., 2011). Further, our insights into the nascent entrepreneurship phenomenon as well as into methodological approaches to study it have grown considerably in the last 15 years. In these respects, Honig and Samuelsson’s (2014) “replication” is very restricted. The dependent variables are partially updated while the rest of the data and the model specification remain the same. If the goal were to settle the issue on the efficacy of business planning for nascent ventures I think a considerably more valuable contribution would have been to use *all* the empirical evidence available today (adding, e.g., PSED; PSED II; CAUSEE) and to address the methods issues of which we are now aware; e.g., the temporality issues discussed above and the endogeneity issue of founders’ self-selecting into business planning activities (Burke et al., 2010).

3. *Methods descriptions are critically important.* Needless to say, methods descriptions need to be detailed and accurate in order to allow *exact* replication. They are also crucial for avoiding misunderstanding and hence misinterpretation of results more broadly. Evidently, aspects of D&S methods description made it difficult for Honig and Samuelsson (2014) to replicate their sample restriction and analysis. It also appears to me that it allowed for a misunderstanding of the extent of problems pertaining to temporality and causal order in D&S analysis. Honig and Samuelsson, in turn, seem to have failed to effectively communicate a key point of their exercise, leading Delmar (2015) to interpret it as an effort to predict the ventures’ status after 5 and 10 years, respectively. Status data cannot possibly have been obtained from all 396 cases in these renewed data collection efforts. As I understand the analysis, what they try to do is to extend the time horizon so as to obtain “final outcomes” (reporting of abandonment or profitability) *at the earliest known point in time* within the studied time span, i.e., an effort to reduce right censoring. However, the method description does not tell for what number of cases data on “final” outcomes was obtained, or when these outcomes were reached. It appears an unspecified number of cases with unknown final outcome were retained in the analysis, with unknown consequences.
4. *Choice of dependent variable.* I would argue that the most important dependent variable in (nascent) entrepreneurship research is “new venture creation success”, i.e., the establishment of a new, viable economic activity or organization (Gartner, 1988; Wiklund et al., 2011). We should acknowledge that indicators of *engagement, persistence or progress* in entrepreneurial processes do not measure venture creation success, whereas H&K’s use of profitability may be an acceptable proxy. Engagement, persistence and progress are likely to have their own antecedents and effects, which can also be theoretically interesting (Davidsson, 2012). But we should call them what they are.
5. *Assessment of “gestation activities”.* The time-stamped “gestation activities” have been one of the most versatile and valuable aspects of the design of PSED-type studies, although opportunities for improvement have also been identified (Davidsson and Gordon, 2012; Gartner et al., 2004). The current debate highlights and clarifies an additional issue: confounding of “what available resources and remaining obstacles pertain to this start-up” and “what have the founders been doing for the specific purpose of getting this venture going”. This goes back to a blurring of levels of analysis that is evident in the very notion of “nascent entrepreneur” (it is actually the venture that is nascent; the founders may be experienced and currently running other ventures). Honig and Samuelsson have reason to be concerned that many founders report activities as completed before the venture is initiated (according to D&S criterion), but their concerns may be exaggerated. As Delmar (2015) alludes to in passing, the result of these activities may be *available to and benefit* the focal, nascent venture although at the time they were not undertaken with that venture in mind. Future projects have reason to consider this difference, and to better capture *on what activities founders spend their time, attention and resources* at different stages of the process (cf. McMullen and Dimov, 2013).
6. *“Null” findings can be important.* The low appreciation of replication is related to the widespread over-reliance on “statistical significance” as a truth criterion (Hubbard and Lindsay, 2013a, 2013b; Schwab et al., 2011; Simmons et al., 2011). Although many individual researchers probably know better, our collective behavior suggests we regard a statistically significant result in a single study (where the assumptions necessary for valid statistical inference were probably already violated) as solid enough evidence to use in the classroom and for policy advice, and to regard subsequent replications as “uninteresting” (usually leading to them not being undertaken). With the focus on statistical significance follows failure to appreciate the importance of *absence* of expected effects (Landis and Rogelberg, 2013). On the substantive issue the core message in H&K and the current extension is that business planning as such does not lead to better performance of nascent ventures. I applaud H&K’s achievement of getting a paper with such a “null finding” as its core message published in a prestigious outlet (and I cannot help wonder whether it would have been possible without the contrast against a “marginally significant” effect on survival to keep reviewers at peace). As the co-author of a recently completed four-year, four-journal journey of the same nature I can testify that convincing colleagues that non-effects can be theoretically and practically important is not an easy task (Davidsson and Gordon, 2015). I also find laudable the current effort to collect and publish additional evidence, stacked against the double institutional norms against “non-significant” results *and* against replication studies. One might wonder what other important, potentially “mythbusting” null findings allude us because they have not found champions that are willing to stand up against the aberrations of our research culture? Again, we have reason to welcome the arrival of JBVi and its broader appreciation of meaningful scholarly contributions.

References

- Aldrich, H.E., Kim, P., Keister, L.A., 2006. Access (not) denied: the effect of financial, human and cultural capital on entrepreneurial entry in the United States. *Small Bus. Econ.* 27 (1), 5–22.
- Bhave, M.P., 1994. A process model of entrepreneurial venture creation. *J. Bus. Ventur.* 9, 223–242.
- Burke, A., Fraser, S., Greene, F.J., 2010. The multiple effects of business planning on new venture performance. *J. Manag. Stud.* 47 (3), 391–415.
- Carter, N.M., Gartner, W.B., Shaver, K.G., Gatewood, E.J., 2003. The career reasons of nascent entrepreneurs. *J. Bus. Ventur.* 18, 13–29.
- Davidsson, P., 2004a. *Researching Entrepreneurship*. Springer, New York.
- Davidsson, P., 2004b. Role models and perceived social support. In: Gartner, W.B., Shaver, K.G., Carter, N.M., Reynolds, P.D. (Eds.), *Handbook of Entrepreneurial Dynamics: The Process of Business Creation*, Sage, Thousand Oaks, pp. 179–185.
- Davidsson, P., 2005. Paul Davidson Reynolds: entrepreneurship research innovator, coordinator and disseminator. *Small Bus. Econ.* 24 (4), 351–358.
- Davidsson, P., 2008. Interpreting performance in entrepreneurship research. In: Davidsson, P. (Ed.), *The Entrepreneurship Research Challenge*, Elgar, Cheltenham, UK.
- Davidsson, P., Gordon, S.R., & Bergmann, H. (Eds.), (2011). *Nascent Entrepreneurship*. 2011a. Cheltenham: Elgar.
- Davidsson, P., 2012. Engagement, persistence, progress and success as theoretically distinct aspects of business creation processes. In: Zacharakis, A., et al. (Eds.), *Frontiers of Entrepreneurship Research 2011*, vol. 31. , Babson College, Wellesley.
- Davidsson, P., Gordon, S.R., 2012. Panel studies of new venture creation: a methods-focused review and suggestions for future research. *Small Bus. Econ.* 39 (4), 853–876.
- Davidsson, P., Gordon, S.R., 2015. Much ado about nothing? The surprising persistence of nascent entrepreneurs through macroeconomic crisis. *Entrep. Theory Practice*. <http://onlinelibrary.wiley.com/doi/10.1111/etap.12152/full>, in press.
- Davidsson, P., Honig, B., 2003. The role of social and human capital among nascent entrepreneurs. *J. Bus. Ventur.* 18 (3), 301–331.
- Davidsson, P., Steffens, P., Gordon, S., 2011b. Comprehensive Australian study of entrepreneurial emergence (CAUSEE): design, data collection and descriptive results. In: Hindle, K., Klyver, K. (Eds.), *Handbook of Research on New Venture Creation*, Elgar, Cheltenham, UK and Northampton, MA, pp. 216–250.
- Davidsson, P., Wiklund, J., 2009. Scott A. Shane: winner of the Global Award for entrepreneurship research. *Small Bus. Econ.* 33 (2), 131–140.
- Delmar, F., 2015. A response to Honig and Samuelsson (2014). *J. Bus. Ventur. Insights* 3, 1–4.
- Delmar, F., Davidsson, P., 2000. Where do they come from? Prevalence and characteristics of nascent entrepreneurs. *Entrep. Reg. Dev.* 12, 1–23.
- Delmar, F., Shane, S., 2003. Does business planning facilitate the development of new ventures? *Strateg. Manag. J.* 24, 1165–1185.
- Gartner, W.B., 1988. "Who is an Entrepreneur?" Is the wrong question. *Am. Small Bus. J.* 12 (4), 11–31.
- Gartner, W.B., Carter, N.M., Reynolds, P.D., 2004. Business start-up activities. In: Gartner, W.B., Shaver, K.G., Carter, N.M., Reynolds, P.D. (Eds.), *Handbook of Entrepreneurial Dynamics: The Process of Business Creation*, Sage, Thousand Oaks, pp. 285–298.
- Hambrick, D.C., 2007. The field of management's devotion to theory: too much of a good thing? *Acad. Manag. J.* 50 (6), 1346–1352.
- Honig, B., Karlsson, T., 2004. Institutional forces and the written business plan. *J. Manag.* 30 (1), 29–48.
- Honig, B., Davidsson, P., Karlsson, T., 2005. Learning strategies of nascent entrepreneurs. *J. Competence-based Manag.* 1 (3), 67–88.
- Honig, B., Samuelsson, M., 2014. Data replication and extension: a study of business planning and venture-level performance. *J. Bus. Ventur. Insights* 1, 18–25.
- Hubbard, R., Lindsay, R.M., 2013a. From significant difference to significant sameness: proposing a paradigm shift in business research. *J. Bus. Res.* 66 (9), 1377–1388.
- Hubbard, R., Lindsay, R.M., 2013b. The significant difference paradigm promotes bad science. *J. Bus. Res.* 66 (9), 1393–1397.
- Kassin, S.M., Dror, I.E., Kukucka, J., 2013. The forensic confirmation bias: problems, perspectives, and proposed solutions. *J. Appl. Res. Mem. Cognit.* 2 (1), 42–52.
- Landis, R.S., Rogelberg, S.G., 2013. Our scholarly practices are derailing our progress: The importance of "nothing" in the organizational sciences. *Ind. Organ. Psychol.* 6 (3), 299–302.
- Liao, J., Welsch, H., Tan, W.L., 2005. Venture gestation paths of nascent entrepreneurs: exploring the temporal patterns. *J. High Technol. Manag. Res.* 16 (1), 1–22.
- McMullen, J.S., Dimov, D., 2013. Time and the entrepreneurial journey: the problems and promise of studying entrepreneurship as a process. *J. Manag. Stud.* 50 (8), 1481–1512.
- O'Boyle, E.H., Banks, G.C., Gonzalez-Mulé, E., 2014. The chrysalis effect: how ugly initial results metamorphosize into beautiful articles. *J. Manag.* <http://dx.doi.org/10.1177/0149206314527133>. (Published online before print).
- Reynolds, P.D., 2007. New firm creation in the US: a PSED overview. *Found. Trends Entrep.* 3 (1), 1–151.
- Reynolds, P.D., Carter, N.M., Gartner, W.B., Greene, P.G., 2004. The prevalence of nascent entrepreneurs in the United States: evidence from the panel study of entrepreneurial dynamics. *Small Bus. Econ.* 23 (4), 263–284.
- Reynolds, P.D., Miller, B., 1992. New firm gestation: conception, birth and implications for research. *J. Bus. Ventur.* 7, 405–417.
- Ruef, M., Aldrich, H.E., Carter, N.M., 2003. The structure of organizational founding teams: homophily, strong ties, and isolation among U.S. entrepreneurs. *Am. Sociol. Rev.* 68 (2), 195–222.
- Samuelsson, M., Davidsson, P., 2009. Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Bus. Econ.* 33 (2), 229–255.
- Schwab, A., Abrahamson, E., Starbuck, W.H., Fidler, F., 2011. Researchers should make thoughtful assessments instead of null-hypothesis significance tests. *Organ. Sci.* 22 (4), 1105–1120.
- Simmons, J.P., Nelson, L.D., Simonsohn, U., 2011. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* 22 (11), 1359–1366.
- Van Gelderen, M., Thurik, A.R., Bosma, N., 2005. Success and risk factors in the pre-startup phase. *Small Bus. Econ.* 24, 365–380.
- Wiklund, J., Davidsson, P., Audretsch, D.B., Karlsson, C., 2011. The future of entrepreneurship research. *Entrep. Theory Pract.* 35 (1), 1–9.