

Online appendix for “The Role of Entrepreneurship in U.S. Job Creation and Economic Dynamism” by Ryan Decker, John Haltiwanger, Ron Jarmin, and Javier Miranda

I. Measurement of Net Growth, Gross Job Creation and Destruction at the Establishment-Level, Firm Level and cell-level.

This section describes the underlying data and the firm-level growth rate measures we use in the paper in more detail. The Longitudinal Business Database (LBD) is the micro data underlying the Business Dynamic Statistics (BDS). We use both the underlying micro data and the public domain data in our analysis. The LBD (Jarmin and Miranda (2002)) covers all business establishments in the U.S. private non-farm sector with at least one paid employee. The LBD begins in 1976 and currently covers over 30 years of data including information on detailed industry and employment for every establishment. We note that the LBD (and in turn the BDS) employment and job creation numbers track closely those of the County Business Patterns (CBP) and Statistics of U.S. Business (SUSB) programs of the U.S. Census Bureau (see Haltiwanger, Jarmin and Miranda (2009)), as they all share the Census Bureau’s Business Register (BR) as their source data. However, due to design features and differences in processing, in particular the correction of longitudinal establishment and firm linkages, the statistics generated from the LBD diverge slightly from those in CBP and SUSB.

The unit of observation in the LBD is the establishment defined as a single physical location where business is conducted. Each establishment-year record in the LBD has a firm identifier associated with it so it is possible to track the ownership structure of firms in any given year as well as changes over time. Firms can own a single establishment or many establishments. In some cases these firms span multiple geographic areas and industries.

Establishments can be acquired, divested, or spun off into new firms so the ownership structure of firms can be very dynamic and complex. We use these firm level identifiers to construct firm level characteristics for each establishment in the LBD.

The construction of firm size measures in the LBD and the BDS is relatively straightforward. Firm size is constructed by aggregating employment across all establishments that belong to the firm. In the LBD and BDS, firm size is available using both the base year and average size methodologies. For base year firm size, we use the firm size for year $t-1$ for all businesses except for new firms. For new firms, we follow the approach used by Birch and others and allocate establishments belonging to firm startups to the firm size class in year t . For average size, we use the average of firm size in year $t-1$ and year t . We use the same approach for new, existing, and exiting firms when using average size. In the paper, we use the average firm size measure (in Figure 3) since as discussed in Haltiwanger, Jarmin and Miranda (2013) this approach to measuring firm size in an analysis of growth dynamics helps take into account bias from the regression to the mean.

To construct firm age measures, we follow the approach adopted for the BDS and based on our prior work (see, e.g., Becker et. al. (2006) and Davis et. al. (2007)). The firm identifiers in the LBD are not explicitly longitudinal. Nevertheless, they are useful for tracking firms and their changing structure over time. A new firm identifier can appear in the LBD either due to a de novo firm birth or due to changes in existing firms. For example, a single location firm opening additional locations is the most common reason for a continuing firm in the LBD to experience a change in firm ID. Other reasons include ownership changes through M&A activity. When a new firm identifier appears in the LBD for whatever reason, we assign the firm an age based upon the age of the oldest establishment that the firm owns in the first year the new

firm ID is observed. The firm is then allowed to age naturally (by one year for each additional year the firm ID is observed in the data) regardless of mergers or acquisitions and as long as the firm ownership and control does not change. An advantage of this approach is that firm births as well as firm deaths are readily and consistently defined. That is, a firm birth is defined as a new firm ID where all the establishments at the firm are new (entering) establishments. Similarly, a firm death is defined as when a firm ID disappears and all of the establishments associated with that firm ID cease operations and exit. If a new firm identifier arises through a merger of two pre-existing firms, we don't treat it as a "firm birth". Rather, the new firm entity associated with the new identifier is given a firm age equal to the age of the oldest continuing establishment of the newly combined entity.

Thus, our firm size and age measures are robust to ownership changes. For a pure ownership change with no change in activity, there will be no spurious changes in firm size or firm age. When there are mergers, acquisitions, or divestitures, firm age will reflect the age of the appropriate components of the firm. Firm size will change but in a manner also consistent with the change in the scope of activity¹.

We now turn to the net growth and job flow measures we use. The firm-level measures used in the paper are constructed from the elemental establishment measures we now describe. Let E_{it} be employment in year t for establishment i . In the LBD, establishment employment is a point-in-time measure reflecting the number of workers on the payroll for the payroll period that includes March 12th. We measure the establishment-level employment growth rate as follows:

$$g_{it} = (E_{it} - E_{it-1}) / X_{it},$$

where

¹ The web appendix to Haltiwanger, Jarmin and Miranda (2013) has a more detailed discussion of this.

$$X_{it} = .5*(E_{it} + E_{it-1}).$$

This growth rate measure has become standard in analysis of establishment and firm dynamics, because it shares some useful properties of log differences but also accommodates entry and exit. (See Davis et al 1996, and Tornqvist, Vartia, and Vartia 1985). The DHS growth rate like the log first difference is a symmetric growth rate measure but has the added advantage that it accommodates entry and exit. It is a second order approximation of the log difference for growth rates around zero. Note that the use of a symmetric growth rate does not obviate the need to be concerned about regression to the mean effects. Also, note the DHS growth rate is not only symmetric but bounded between -2 (exit) and 2 (entrant).

Note that the DHS growth rate measure can be flexibly defined for different aggregations of establishments including firms. We first discuss the measures of net growth used in the analysis. In particular, consider the following relationships

$$g_t = \sum_s (X_{st} / X_t) g_{st} = \sum_s ((X_{st} / X_t) \sum_{i \in s} (X_{it} / X_{st}) g_{it})$$

where

$$X_t = \sum_s X_{st} = \sum_s \sum_{i \in s} X_{it}$$

where g_t is the aggregate DHS growth rate and s indexes classifications of establishments into groups defined for any level of aggregation s where s can refer to firm, industry, firm size, or firm age classifications. Thus, the DHS net growth rates for various aggregations of interest are just properly weighted averages of establishment-level growth rates where the establishment is the lowest level of aggregation in the LBD. Important groupings for this paper include firms and firm size and age categories.

These aggregation properties of DHS growth rates imply that the aggregate growth rate (or the growth rate for any cell) can be equivalently measured from the employment-weighted average of the growth rates of business units or from aggregate level data directly:

$$g_{st} = \sum_{e \in s} (X_{et} / X_{st}) g_{et} = (E_{st} - E_{s,t-1}) / X_{st}$$

As discussed in the text, we also compute firm-level growth rates in a manner that captures organic growth and not that due to M&A activity. We do this as follows. The period $t-1$ to period t net growth rate for a firm is the sum of the appropriately weighted DHS net growth rate of all establishments owned by the firm in period t , including acquisitions, plus the net growth attributed to establishments owned by the firm in period $t-1$ that it has closed before period t . For any continuing establishment that changes ownership, this method attributes any net employment growth to the acquiring firm. Note, however, if the acquired establishment exhibits no change in employment, there will be no accompanying change in firm level employment induced by this ownership change.

We use the establishment and firm-level growth rate measures to compute not only net growth but also job creation and job destruction (and the related job creation from entry and job destruction from exit). Measures of job creation and destruction at the establishment level are given by:

$$JC_{it} = \max(g_{it}, 0)$$

$$JD_{it} = \max(-g_{it}, 0)$$

Job creation from entry at the establishment level is given by:

$$JC_{it} = \max(g_{it}, 0) * I\{g_{it} = 2\}$$

where I is an indicator variable equal to one if expression in brackets hold, zero otherwise, and $g_{it} = 2$ denotes an entrant. Similarly job destruction from exit at the establishment level is given by:

$$JD_{it} = \max(-g_{it}, 0) * I\{-g_{it} = 2\}$$

where $g_{it} = -2$ denotes an exit.

Using these measures it is straightforward to generate aggregate measures of job creation and destruction as well as job creation and destruction from entry and exit, respectively (at any level of aggregation) given by:

$$JC_t = \sum_i (X_{it} / X_t) \max\{g_{it}, 0\}$$

$$JD_t = \sum_i (X_{it} / X_t) \max\{-g_{it}, 0\}$$

$$JC_Cont_t = \sum_{i \in c} (X_{it} / X_t) \max\{g_{it}, 0\}$$

$$JD_Cont_t = \sum_{i \in c} (X_{it} / X_t) \max\{-g_{it}, 0\}$$

$$JC_Entry_t = \sum_i (X_{it} / X_t) I\{g_{it} = 2\} \max(g_{it}, 0).$$

$$JD_Exit_t = \sum_i (X_{it} / X_t) I\{g_{it} = -2\} \max(-g_{it}, 0)$$

Given these definitions, the following simple relationships hold:

$$g_t = JC_t - JD_t, \quad JC_t = (X_{ct} / X_t) JC_Cont_t + JC_Entry_t \quad \text{and}$$

$$JD_t = (X_{ct} / X_t) JD_Cont_t + JD_Exit_t$$

where JC_Cont and JD_Cont are job creation and job destruction for continuing establishments respectively. Note they are defined here only over the range of continuing establishments so that the share of employment accounted for by continuing establishments needs to be used in aggregation.

We can also define job creation and destruction starting from the firm level in a manner analogous to the relationships above. As with net changes in the presence of M&A activity,

appropriate care must be used in considering the relationship between establishment-level and firm-level employment gross rate patterns. But even after focusing on only organic growth, there are inherent differences in firm and establishment-level gross rates of creation and destruction. Firms that are expanding (including firm entrants) contribute to firm-level gross job creation and firms that are contracting (including firm exits) contribute to firm-level gross job destruction. But a firm that is expanding may have some contracting (or exiting) establishments and a firm that is contracting may have some expanding (or entering) establishments. As such, gross job creation and destruction is inherently greater when summed up from establishment-level net changes than when summed up from firm-level net changes. In a closely related way, we note that firm entry implies that there must be establishment entry as an entering firm is identified as a new firm with all new establishments but the converse does not hold. Similarly, firm exit implies that there must be establishment exit but the converse does not hold.

The establishment-level and firm-level data can be aggregated to the cell-level “s” for any definition of cell “s”. The following relationships hold:

$$g_{st} = JC_{st} - JD_{st}, \quad JC_{st} = (X_{scf} / X_{st})JC_{Cont_{st}} + JC_{Entry_{st}} \quad \text{and}$$

$$JD_{st} = (X_{cst} / X_{st})JD_{Cont_{st}} + JD_{Exit_{st}}$$

These relationships imply that one can decompose overall net growth into its continuer, job creation from entry and job destruction from exit components at the cell-level:

$$g_{st} = (X_{st} / X_t)g_{cst} + JC_{Entry_{st}} - JD_{Exit_{st}}$$

Summary measures of reallocation are readily constructed from these measures. We follow Davis, Haltiwanger and Schuh (1996) here as well. Job Reallocation for cell s at time t is given by:

$$JR_{st} = JC_{st} + JD_{st}$$

A measure of reallocation that abstracts from the net changes in the cell s at time t is what has been called excess job reallocation:

$$Excess_JR_{st} = JC_{st} + JD_{st} - abs(JC_t - JD_{st})$$

II. TFPR vs. TFPQ Measures Used in the Literature on Productivity and Firm Dynamics

For the discussion on startups and productivity growth we rely on the findings in the literature. It is helpful to provide some background discussion on the measures that are used in the literature. The studies we cite use a TFP index at the plant-level similar to that used in Baily, Hulten and Campbell (1992) and a series of papers that built on that work (see Syverson (2011) for an excellent survey). The index is given by:

$$\ln TFP_{et} = \ln Q_{et} - \alpha_K \ln K_{et} - \alpha_L \ln L_{et} - \alpha_M \ln M_{et} \quad (1)$$

where Q is real output, K is real capital, L is labor input, M is materials, α denotes factor elasticities, the subscript e denotes individual establishments and the subscript t denotes time. A key issue is how real output is measured. In most of the literature, nominal output is measured as total shipments plus the change in inventories. Real output is measured by deflating nominal output using an industry-level measure. In this respect, almost all studies we cite use what Foster, Haltiwanger and Syverson (2008) call revenue productivity or TFPR. That is, measured productivity captures plant-specific variation in output prices which may reflect a variety of factors (e.g., demand, costs, quality and technical efficiency itself).

As discussed in the text of the paper, a small number of studies have contrasted the patterns in the micro data using a measure of productivity that abstracts from plant-level price variation. For selected products in the U.S., physical quantities are available so that a measure of

output based on such physical output can be constructed with an accompanying measure of physical productivity (called TFPQ). Foster, Haltiwanger and Syverson (2008) construct and analyze the relationship between TFPQ and TFPR for 11 such products. Eslava et. al. (2013) construct and analyze the relationship between TFPQ and TFPR for a much wider set of industries for Colombia since in that country the underlying Census of Manufactures (conducted annually) is collected in an integrated fashion with the Producer Price Index data. This data integration permits the construction of plant-specific price indices for measuring real output.

Similar patterns are found in both the U.S. and Colombia (all of these statistics reflect variation within detailed industries). First, TFPQ and TFPR are highly correlated (0.75 in the U.S., 0.70 in Colombia). TFPQ is inversely correlated with prices (-0.54 in the U.S., -0.65 in Colombia) which is consistent with models of product differentiation: that is, higher-productivity plants have lower marginal costs and charge lower prices. Moreover, there is a moderate but positive covariance between firm size as measured by physical output and productivity in the data (0.28 in U.S.).

The relationship between these alternative measures of productivity and growth and survival are similar. This robustness can be illustrated by comparing the estimated marginal effects of productivity on the probability of exit across studies and samples. In the Foster, Grim and Haltiwanger (2013) estimates using pooled data for all U.S. manufacturing industries the estimated marginal effect of revenue productivity (TFPR) on exit is -0.060 (0.003). The estimate from Foster, Haltiwanger and Syverson (2008) using selected industries (with physical quantity data) for the equivalent marginal effect of revenue productivity (TFPR) on exit is -0.063 (0.014). When revenue productivity is decomposed into physical productivity and price components and both components are included in the same empirical model of selection the

estimated marginal effects of physical productivity (TFPQ) on exit is -0.062 (0.014) and the estimated marginal effect of prices on exit is -0.069 (0.021).

References

- Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. Productivity Dynamics in Manufacturing Plants. In *Brookings Papers on Economic Activity: Microeconomics*, ed. Clifford Winston and Martin Neil Baily. Washington, DC: Brookings Institution Press.
- Becker. Randy A., John Haltiwanger, Ron S. Jarmin, Shawn D. Klimek and Daniel J. Wilson, 2006, "Micro and Macro Data Integration: the Case of Capital", in *A New Architecture for the U.S. National Accounts*, (Jorgenson, Landefeld and Nordhaus eds), NBER/University of Chicago Press.
- Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda. 2007. "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms." Chap. 2 in *NBER Macroeconomics Annual 2006* edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, vol. 21, pp. 107-80. Cambridge, MA: MIT Press.
- Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. Cambridge, MA: MIT Press.
- Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler. 2004. "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia." *Journal of Development Economics* 75(2): 333-72.
- Eslava, Marcela and John Haltiwanger. 2013. "Trade Reforms and Market Selection: Evidence from Manufacturing Plants in Colombia." *Review of Economic Dynamics* 16(1): 135-58.

Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2013. "Reallocation in the Great Recession: Cleansing or Not?" CES Working Paper no. 13-42. Washington, D.C.: Center for Economic Studies, U.S. Census Bureau.

Foster, Lucia, John Haltiwanger, and Chad Syverson, 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1), 394-425.

Haltiwanger, John, Jarmin, Ron, and Javier Miranda. 2009. "Business Dynamics Statistics: An Overview." Kauffman Foundation: Other Research. At http://www.kauffman.org/uploadedFiles/BDS_handout_011209.pdf

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. 2013. "Who Creates Jobs? Small vs. Large vs. Young." *Review of Economics and Statistics* 95(2), pp. 347-361.

Jarmin, Ron S., and Javier Miranda. 2002. "The Longitudinal Business Database." CES Working Paper no. 02-17. Washington, D.C.: Center for Economic Studies, U.S. Census Bureau.

Syverson, Chad, 2004. "Substitutability and Product Dispersion," *Review of Economics and Statistics*, 86 (2), 534-550.

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49(2): 326-65.

Törnqvist, Leo, Pentti Vartia, and Yrjö O. Vartia. 1985. "How Should Relative Changes be Measured?" *The American Statistician* 39(1): 43-6.