

Unpublished Appendix for
TAXES AND GROWTH IN A FINANCIALLY
UNDERDEVELOPED COUNTRY: EVIDENCE FROM
THE CHILEAN INVESTMENT BOOM

Chang-Tai Hsieh

University of California Berkeley and NBER

Jonathan A. Parker

Princeton University and NBER

June 2007

Appendix

A Additional details on the Chilean manufacturing census

The *CMC* survey questionnaires do not directly ask for asset information, but according to the instructions of the survey, plants are supposed to provide a copy of their balance sheets (that they are required to keep for tax purposes). These sheets are the source of the asset information, including the book value of capital, contained in the Survey in 1980 and 1981. Book value is also asked annually since 1992. There is a “book value of assets” measure reported in 1986 and 1987. The 1986 and 1987 numbers are implausibly small (i.e. less than 5 percent of investment) and we do not use them.

We exclude buildings from our measure of capital and investment because the book value of capital in 1980 and 1981 only lists land and buildings together while investment data until 1987

do not include land at all. That is, in 1980 or 1981 we construct capital stock using reports on book value of fixed assets in machinery and equipment, furniture, and vehicles and then distribute the amount reported in “other fixed assets” across the three categories (machinery, vehicles and real estate) in the proportion that each category is of the sum of the three. From 1992 on, the reported asset types are: land; buildings; machinery and equipment; and vehicles. The *CMC* contains information on four main categories of capital investment: buildings; machinery and equipment; vehicles; and, after 1986, land.¹

In the World Bank extract, the reported book value of capital stock has a monetary adjustment factor (in addition to being deflated to be made real) and is also adjusted to account for some depreciation. Since the documentation is unclear as to how these adjustments are derived, we construct two separate data extracts: one that is based on the reported book value of capital as reported in the raw survey data for 1980 (1981 if 1980 is missing or zero); the other that is based on the “net,” “inflation adjusted” capital stock as reported in the World Bank extract for 1980 (1981 if 1980 is missing or zero). “Net” and “inflation adjusted” are the terms used in the World Bank documentation.

We use a machinery price index to deflate both investment and fixed assets in machinery and equipment, and vehicles. Capital stock in “other assets” is distributed across categories in the construction of K by category and the deflation is done after this distribution. We deflate the reported book values of capital stock by the average deflator for year t and year $t + 1$ since the reported book values refer to end of year values and the deflators provide a price index for the entire year. Flow variables such as investment are simply deflated by year t price deflators.

In cumulating past capital stocks and investment to construct capital stocks, the data are cleaned in three ways. First, in some rare cases, we infer that the capital stock becomes negative, and we reset these stocks to zero. Since capital is a denominator such an observation is not used in analysis. Since we drop extreme outliers some cases close in time to this observation with unreasonably low capital stocks are also not used. Second, when a plant disappears from the sample for only one year, we assume that it was merely missed in the survey for a year and

¹In 1987 and 1988 investment in land is reported under “other” investment; after 1988, land is its own investment category. Prior to 1987, investment in land is not included in the survey; notably it is not included in investment in real estate (whereas in this period land is included in capital stock in real estate).

carry its capital stock forward over the missing year without adjustment. That is we assume that investment roughly equalled depreciation during the omitted year. Finally, plants absent for more than one year are considered to have become too small to be in the survey or to have gone bankrupt, and any future observations on such a plant are dropped.

A nontrivial number of plants, about 20 percent, switch industry codes at some point during the period over which they are observed. We treat these as legitimate. Firms that are associated with plants that cover multiple industries are associated with the mode industry. Where ties occur they are broken somewhat randomly by choosing the smallest industry code. Most firms operate plants in related industries, and we verified that this allocation rule does not drive any results. Finally, roughly half of plants “attrits” between 1983 and 1990 for the alternative sample that sets investment to missing following two consecutive years of zero reported investment.

Industries are defined by their three-digit ISIC code, except that a) the food production industry (311) is treated as four separate three digit industries since it has so many of the firms in the sample (groups are 3111 and 3112 (meat and dairy); 3113, 3114, and 3115 (canning and oils and fats); 3116, 3118 and 3119 (grains, sugar, and cocoa); 3117 (bakeries); b) 312 (other manufactured foods and animal feed) is grouped with 3140 (tobacco); c) industry 381 (fabricated metal except machines) is also treated more finely – it is simply left at the four digit level; and finally d) due to small numbers, industry 3540 (petroleum and coal) is grouped with 353 (petroleum refineries) and 3620 (glass) is grouped with 361 (pottery and china).

The trade data was provided by Nina Pavcnik. In her data, each 4-digit industry in each year is considered a net exporter, a net importer, or a non-trading industry. (While there is also more detailed information on ratios of exports and imports to output, in personal communication along with the data, Nina Pavcnik recommended using only the classification into these three categories). We assign values of 1, -1, and 0 to each status respectively and then average across plants in a three-digit industry in each year to assign a numerical value to each three digit industry in each year. There is no trade data for ciu’s 3131-3134, 3721-3729 and 3903 and 3909. We use the average of the 1980 and 1981 values (interacted with year dummies) for each industry in our regressions.

B Additional data sources

Figure 1a: Penn-World Tables.

Figures 1b, 3a, and 5: Bennett, Schmitt-Hebbel, and Soto (1999).

Figure 3b: International Financial Statistics, IMF.

Figure 4a: Central Bank of Chile, "Quarterly Economic and Financial Report of Chile," various issues.

Figure 4b: Ministerio de Hacienda, Direccion de Presupuestos, "Estadisticas de las Finanzas Publicas, 1988-1997," March 1998 and Ministerio de Hacienda, Direccion de Presupuestos, "Ley de Presupuestos del Sector Publico," various years.

Figures 6a and 6b: Eyzaguirre and Lefort (1999), Tables 3-1 and 3-6.