

# Online Appendix: Entry into Export Markets

## 1 Data

All the firms in our sample have Chile as country of origin and we restrict our analysis to firms whose main activity is the manufacture of chemicals and chemical products (ISIC sector number 24). This is the second largest export manufacturing sector in Chile. Our data set includes both exporters and non-exporters from 1995 to 2005. However, we restrict our sample to firms that have been active in Chile for all eleven years between 1995 and 2005.

Our data come from two separate sources. The first is an extract of the Chilean customs database, which covers the universe of exports of Chilean firms. The second is the Chilean Annual Industrial Survey (*Encuesta Nacional Industrial Anual*, or ENIA), which includes all manufacturing plants with at least ten workers. We merge these two data sets using firm identifiers, allowing us to exploit information on the export destinations of each firm and on their domestic activity. These two datasets provide us with information on the value exported per firm and destination country and the firm characteristics that we use to predict the actual revenue from exporting.<sup>1</sup>

We complement our customs-ENIA data with a database of country characteristics from CEPIL.<sup>2</sup> This database gives us information on the physical distance between any country and Chile, as well as on the gravity variables that we include in  $X_{ict}^R$ . Finally, we collect data on real GDP per capita from the World Bank World Development Indicators.

## 2 Revenue Regression

In the revenue regression, the observed right hand side covariate vector,  $X_{ict}^R$ , includes:<sup>3</sup> (a) firm characteristics: number of skilled and unskilled workers ( $Sk.Emp$  and  $UnSk.Emp$ ), average wage of skilled and unskilled workers ( $Sk.Wage$  and  $UnSk.Wage$ ), and average value added per worker ( $VA/Emp$ ); (b) country characteristics: dummy for being landlocked ( $Landlocked$ ), country-specific level of imports in chemical sector ( $Sec.Imp.$ ), and GDP per capita ( $GDPpc$ ); (c) gravity variables or distance measures: physical distance between Chile and destination country ( $Dist.$ ), dummy for having a FTA with Chile ( $FTA$ ), and dummies for sharing border ( $Border$ ), continent ( $South America$ ), language ( $Spanish$ ), and World Bank GDP per capita group with Chile ( $Med-High Inc.$ ). In addition, we introduce firm and year fixed effects, and allow a dummy capturing that country  $c$  is equal to Chile to change by year. We check the robustness of our results to outliers by running three separate regressions: regression *I* includes all the observed export flows; regression *II* drops the largest 3%; and regression *III* drops both the smallest and largest 3%.

We estimate the parameter vector  $\theta$  using Nonlinear Least Squares (NLLS). The estimates of  $\theta$  are contained in the following table.

---

<sup>1</sup>An observation in this data is a firm-country-year combination. For each observation we have information on the value of goods sold in US dollars. We obtain sales values in year 2000 terms using the US CPI.

<sup>2</sup>Available at <http://www.cepii.fr/anglaisgraph/bdd/distances.htm>. Mayer and Zignago (2006) provide a detailed explanation of the content of this database.

<sup>3</sup>We write in parenthesis the label that each variable has in column 1 of Table 1.

Table 1: Revenue Regression

	<i>I</i>	<i>II</i>	<i>III</i>
Variables	$R_{ict}$	$R_{ict}$	$R_{ict}$
Chile	0.603 (0.371)	-0.841 <sup>a</sup> (-2.322)	-0.858 <sup>a</sup> (-2.337)
ln(Emp).*Chile	0.247 <sup>a</sup> (7.285)	0.442 <sup>a</sup> (11.623)	0.440 <sup>a</sup> (11.418)
ln(Sk.Emp.)	0.151 <sup>a</sup> (7.393)	0.0315 (1.232)	0.032 (1.241)
ln(UnSk.Emp.)	0.175 <sup>a</sup> (16.979)	0.026 <sup>a</sup> (2.554)	0.026 <sup>a</sup> (2.539)
ln(Sk.Wage)	0.163 <sup>a</sup> (13.208)	0.107 <sup>a</sup> (6.808)	0.107 <sup>a</sup> (6.698)
ln(UnSk.Emp.)	-0.033 <sup>a</sup> (-6.829)	-0.003 (-0.738)	-0.003 (-0.711)
ln(VA/Emp.)	0.418 <sup>a</sup> (51.229)	0.389 <sup>a</sup> (24.959)	0.388 <sup>a</sup> (24.549)
ln(Dist.)	-0.772 <sup>a</sup> (-1.871)	-1.429 <sup>a</sup> (-17.927)	-1.427 <sup>a</sup> (-17.656)
Border	-2.630 <sup>a</sup> (-3.295)	-1.961 <sup>a</sup> (-20.693)	-1.957 <sup>a</sup> (-20.376)
Spanish	0.090 (0.312)	0.712 <sup>a</sup> (14.868)	0.711 <sup>a</sup> (14.617)
South America	-0.225 (-0.403)	-0.047 (-0.462)	-0.046 (-0.451)
Med-High Inc.	-0.102 (-0.280)	-0.434 <sup>a</sup> (-6.387)	-0.433 <sup>a</sup> (-6.279)
FTA	-0.071 (-0.305)	0.171 <sup>a</sup> (4.276)	0.170 <sup>a</sup> (4.193)
Landlocked	-1.396 (-0.429)	-2.722 <sup>b</sup> (-1.778)	-2.709 <sup>b</sup> (-1.768)
Sec.Imp.	0.987 <sup>a</sup> (6.561)	1.259 <sup>a</sup> (39.484)	1.257 <sup>a</sup> (38.885)
ln(GDPpc)	-0.473 <sup>a</sup> (-4.179)	-0.367 <sup>a</sup> (-16.199)	-0.366 <sup>a</sup> (-15.929)
Firm F.E.	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes
(Year F.E.)*Exp	Yes	Yes	Yes
<i>Adj.R</i> <sup>2</sup>	0.990	0.907	0.906
Left Trunc.	0%	0%	3%
Right Trunc.	0%	97%	97%
Obs.	6931	6724	6516

In order to estimate the numbers in Table 1, we exclusively use data for those observations with positive exports. Therefore, the estimated vector  $\hat{\theta}$  verifies the following mean independence condition

$$\mathbb{E}[R_{ict} - r(X_{ict}^R; \hat{\theta}) | X_{ict}^R, \mathbf{d}_{ict} = 1] = \mathbb{E}[\nu_{ict}^R + \varepsilon_{ict}^R + \eta_{ict}^R | X_{ict}^R, \mathbf{d}_{ict} = 1] = 0,$$

where  $\eta_{ict}^R = r(X_{ict}^R; \theta) - r(X_{ict}^R; \hat{\theta})$ . Given that  $\mathbf{d}_{ict}$  is potentially a function of  $\nu_{ict}^R$ , the estimate  $\hat{\theta}$  will generally not be a consistent estimate of  $\theta$ . More precisely,  $\hat{\theta}$  is a consistent estimator of  $\theta$  if and only if

$$\mathbb{E}[R_{ict} - r(X_{ict}^R; \theta) | X_{ict}^R, \mathbf{d}_{ict} = 1] = 0, \quad (1)$$

This moment condition will hold if (a)  $\nu_{ict}^R = 0$ ; and, (b)  $\mathbb{E}[\varepsilon_{ict}^R | X_{ict}^R] = 0$ .<sup>4</sup>

Even in those cases in which  $\hat{\theta}$  is not a consistent estimator of  $\theta$ , we can still consider  $r(X_{ict}^R; \hat{\theta})$  as a reduced form approximation of the potential revenue from exporting,  $R_{ict}$ , and, moreover, as a proxy for the expected revenue from exporting,  $R_{ict}^*$ . The resulting error term from using  $\hat{R}_{ict} = r(X_{ict}^R; \hat{\theta})$  as a proxy for  $R_{ict}^*$  is

$$R_{ict}^* - \hat{R}_{ict} = R_{ict}^* - r(X_{ict}^R; \hat{\theta}) = R_{ict}^* - R_{ict} + R_{ict} - r(X_{ict}^R; \hat{\theta}),$$

and it holds that  $\mathbb{E}[R_{ict}^* - R_{ict} | \mathcal{J}_{it}] = 0$  and  $\mathbb{E}[R_{ict} - r(X_{ict}^R; \hat{\theta}) | X_{ict}^R, \mathbf{d}_{ict} = 1] = 0$ . The first orthogonality condition arises from the definition of  $R_{ict}^*$  as the expected revenue conditional on the information set  $\mathcal{J}_{it}$ . The second orthogonality condition is a direct implication of the definition of  $\hat{\theta}$  as the result of projecting realized revenues  $R_{ict}$  on the nonlinear function  $r(X_{ict}^R; \hat{\theta})$ .

### 3 Instrument Functions for Empirical Exercise

We present the results for three different sets of instruments. Remember from the main text that the vector  $X_{ict}$  contains the proxy  $\hat{R}_{ict}$  that we use for the unobserved expected revenue  $R_{ict}^*$ . The vector  $Z_{1ict}$  contains the distance covariate  $D_c$ , measured without error, and the variable  $Z_{2ict}$  incorporates the lagged revenue  $\hat{R}_{ict-1}$  as an instrument for current revenue. The vector  $Z_{ict}$  is defined as  $(Z_{1ict}, Z_{2ict})$ .

The *first set of instruments* incorporates the following six instruments:

$$\Psi_r^{(1)}(X_{ict}, Z_{1ict}) = \begin{bmatrix} \begin{pmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{pmatrix} \\ - \begin{pmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{pmatrix} \end{bmatrix}, \quad \Psi_s^{(1)}(X_{ict}, Z_{1ict}) = \begin{bmatrix} \begin{pmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{pmatrix} \lambda(X_{ict}, Z_{1ict}) \\ - \begin{pmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{pmatrix} \lambda(X_{ict}, Z_{1ict}) \end{bmatrix} \quad (2)$$

<sup>4</sup>A sufficient condition for restriction (b) is that  $X_{ict}^R \in \mathcal{J}_{it}$ . Note that we have not imposed this restriction so far and that it is not necessary;  $X_{ict}^R$  could be mean independent of  $\varepsilon_{ict}^R$  even if  $X_{ict}^R \notin \mathcal{J}_{it}$ .

where  $\Psi_r^{(1)}(\cdot)$  is the set of instruments that we apply to the revealed preference inequalities,  $\Psi_s^{(1)}(\cdot)$  is the set applied to the score function inequalities,

$$\lambda(X_{ict}, Z_{1ict}) = \frac{f_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))},$$

and  $u_{ict}(X_{ict}, Z_{1ict}; \beta) = \beta_1 \hat{R}_{ict} - \beta_2 - \beta_3 D_c$ . The set of instruments in equation (2) defines inequalities that contain the true value of the parameter vector  $\beta$  only under the assumption of no expectational error (i.e. perfect foresight). This set of instruments  $\Psi^{(1)}(\cdot)$  generates identified parameters that are identical to those arising from maximizing the log likelihood function:

$$L = \mathbb{E}[L(X_{ict}, Z_{1ict}, \mathbf{d}_{ict})] = \mathbb{E}[\mathbf{d}_{ict} \log(1 - F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))) + (1 - \mathbf{d}_{ict}) \log(F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta)))] \quad (3)$$

In order to prove that our moment inequalities, combined with the instrument vector in equation (2), identify the same parameter value as the maximum likelihood estimator, we show that these moment inequalities are equivalent to the first order condition from the maximization of the log likelihood function. In particular, the vectoral first order condition from maximizing the log likelihood function in equation (3) is:

$$\begin{aligned} & \mathbb{E} \left( \begin{bmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{bmatrix} \cdot \left[ \mathbf{d}_{ict} \frac{f_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{1 - F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} - (1 - \mathbf{d}_{ict}) \frac{f_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} \right] \right) = \\ & \mathbb{E} \left( \begin{bmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{bmatrix} \cdot \left[ \frac{f_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} \right] \cdot \left[ \mathbf{d}_{ict} \frac{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{1 - F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} - (1 - \mathbf{d}_{ict}) \right] \right) = 0 \end{aligned}$$

and this first order condition is equivalent to the combination of the following two inequalities:

$$\begin{aligned} & \mathbb{E} \left( \begin{bmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{bmatrix} \cdot \left[ \frac{f_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} \right] \cdot \left[ \mathbf{d}_{ict} \frac{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{1 - F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} - (1 - \mathbf{d}_{ict}) \right] \right) \geq 0, \\ & -\mathbb{E} \left( \begin{bmatrix} \hat{R}_{ict} \\ 1 \\ D_c \end{bmatrix} \cdot \left[ \frac{f_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} \right] \cdot \left[ \mathbf{d}_{ict} \frac{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{1 - F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} - (1 - \mathbf{d}_{ict}) \right] \right) \geq 0, \end{aligned}$$

We can rewrite these two inequalities as

$$\mathbb{E} \left( \Psi_s^{(1)}(X_{ict}, Z_{1ict}) \left[ \mathbf{d}_{ict} \frac{F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))}{1 - F_\nu(-u_{ict}(X_{ict}, Z_{1ict}; \beta))} - (1 - \mathbf{d}_{ict}) \right] \right) \geq 0,$$

where  $\Psi_s^{(1)}(X_{ict}, Z_{1ict})$  is defined in equation (2).

The *second set of instruments* is:

$$\Psi_r^{(2)}(Z_{ict}) = \begin{bmatrix} \begin{pmatrix} \hat{R}_{ict-1} \\ 1 \\ D_c \\ 1/D_c \\ 1/\sqrt{D_c} \end{pmatrix} \\ - \begin{pmatrix} \hat{R}_{ict-1} \\ 1 \\ D_c \\ 1/D_c \\ 1/\sqrt{D_c} \end{pmatrix} \end{bmatrix}, \quad \Psi_s^{(2)}(Z_{ict}) = \begin{bmatrix} \begin{pmatrix} \hat{R}_{ict-1} \\ 1 \\ D_c \\ 1/D_c \\ 1/\sqrt{D_c} \end{pmatrix} \cdot \lambda(Z_{ict}) \\ - \begin{pmatrix} \hat{R}_{ict-1} \\ 1 \\ D_c \\ 1/D_c \\ 1/\sqrt{D_c} \end{pmatrix} \cdot \lambda(Z_{ict}) \end{bmatrix}, \quad (4)$$

where  $\Psi_r^{(2)}(\cdot)$  is the set of instruments that we apply to the revealed preference inequalities,  $\Psi_s^{(2)}(\cdot)$  is the set applied to the score function inequalities,

$$\lambda(Z_{ict}) = \frac{f_\nu(-u_{ict}(Z_{ict}; \beta))}{F_\nu(-u_{ict}(Z_{ict}; \beta))},$$

and  $u_{ict}(Z_{ict}; \beta) = \beta_1 \hat{R}_{ict-1} - \beta_2 - \beta_3 D_c$ .

The *third set of instruments* simply expands the second set to incorporate interaction terms of lagged revenues and distance:

$$\Psi_r^{(3)}(Z_{ict}) = \begin{bmatrix} \Psi_s^{(2)}(Z_{ict}) \\ \begin{pmatrix} \hat{R}_{ict-1}(1/D_c) \\ \hat{R}_{ict-1}(1/\sqrt{D_c}) \end{pmatrix} \\ - \begin{pmatrix} \hat{R}_{ict-1}(1/D_c) \\ \hat{R}_{ict-1}(1/\sqrt{D_c}) \end{pmatrix} \end{bmatrix}, \quad \Psi_s^{(3)}(Z_{ict}) = \begin{bmatrix} \Psi_s^{(2)}(Z_{ict}) \\ \begin{pmatrix} \hat{R}_{ict-1}(1/D_c) \\ \hat{R}_{ict-1}(1/\sqrt{D_c}) \end{pmatrix} \cdot \lambda(Z_{ict}) \\ - \begin{pmatrix} \hat{R}_{ict-1}(1/D_c) \\ \hat{R}_{ict-1}(1/\sqrt{D_c}) \end{pmatrix} \cdot \lambda(Z_{ict}) \end{bmatrix}. \quad (5)$$

Both the second and third set of instruments are a function of  $Z_{2ict} = \hat{R}_{ict-1}$  (instead of  $X_{ict} = \hat{R}_{ict}$ ). Therefore, they are both consistent with the presence of expectational error and will generate identified sets that will contain the true value of the parameter vector even in the presence of this additional source of error.

## References

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's Distances Measures," *mimeo*, May 2006.