Local Export Spillovers in an Developing Country*

Jorge F. Chavez†

June 1, 2016

Work in progress. Please do not circulate.

Abstract

This paper investigates the importance of export spillovers for the performance of new export flows initiated by Peruvian manufacturing firms. Using export transaction level data as well as data on the location of plants of Peruvian manufacturers, I find evidence suggesting that the proximity to other exporting manufacturers has non-negligible effects on the probability of exporting a new product-country combination (extensive margin) and on the probability of survival of this new trade flow (a less explored dimension of the intensive margin). This evidence is consistent with the predictions of a simple theoretical framework in which proximity to other exporters affects both margins of new firm-level export flows as a result of uncertainty surrounding fixed costs of exporting. Results are significant and robust to a variety of checks. From a policy point of view the conclusions from this paper can be relevant for the design of both industrial policies in general (e.g. the relevance of supporting the creation of industrial parks), and for the design of export promotion policies oriented at facilitating the successful internationalization of domestic firms in particular.

JEL Classification: F1, R12

Keywords: Export spillovers, transaction-level export data, extensive margin, survival.

*I thank Rocco Macchiavello, Dennis Novy, Guillermo Noguera, Christian Volpe and Alejandro Riano and for insightful comments on early drafts of this paper. I also thank to participants at the 2015 European Trade Study Group in Paris, the Applied Micro Early Stage Group and the CAGE Work in Progress Group both at Warwick. I am also grateful to Sociedad de Comercio Exterior del Perú (COMEX-Peru) for facilitating access to the data used in this paper. The usual disclaimers apply.

†University of Warwick, Department of Economics. E-mail: j.chavez-cotrado@warwick.ac.uk.

1 Introduction

There is a large consensus in the international trade literature that entry into new export markets is both costly (Roberts and Tybout 1995, Clerides et al. 1998, Bernard and Jensen 2004; Melitz 2003, Das et al. 2007) and associated with high levels of uncertainty (Rauch and Watson 2003; Albornoz-Crespo et al. 2010; Aeberhardt et al. 2009; Segura-Cayuela and Vilarrubia 2008; Eaton et al. 2012). One consequence of these factors is the empirically established fact that exporters tend to start selling small amounts and facing low probabilities of survival (Eaton et al. 2007; Freund and Pierola 2010; Buono et al. 2008; Amador and Opromolla 2008).

In this line, much of the emphasis in the literature has been in uncovering a number of factors that affect the export performance of firms, including total factor productivity, age and experience. Now we know that only the most productive firms are able to engage in international trade, that firms tend to be more successful if they become exporters late in their life cycle, and that their expansion across destination countries and products is sequential and is closely linked to their experience in similar markets. However, there are still other less explored elements that may affect the behavior and performance of exporting firms. One of these factors are the external economies from the concentration of exporters in a given geographical space, an effect that has been termed *export spillovers* (see Aitken et al. 1997, Koenig et al. 2010). In particular, the location of a firm is a key determinant for its success in export markets not only because it determines access to input sources and to key infrastructure services, but also because the presence of neighbors selling to similar export markets can have positive effects in shaping the firm's export performance.

Proximity to other exporters can facilitate the access to foreign markets through the interaction of both positive and negative externalities which are components of the broader agglomeration economies classified by Duranton and Puga (2004) into sharing, matching and learning mechanisms. First, proximity to other exporters allow the development and *sharing* of specialized local input markets and intermediate industries such as packing products, transportation services or cold-storages, as well as infrastructure and local public goods that contribute to enhance the productivity of all firms. Second, *matching* refers to the quality of matches between workers and firms on the labor market, as well as local buyer-supplier networks of intermediate products. Denser areas are expected to provide higher-quality matches because of the larger number of opportunities that agents face in a thicker market. Hiring of managers or workers with relevant past experience in the production of goods for a particular market is also more likely when firms are located near other exporters (see for example Choquette and Meinen 2014). Finally, proximity facilitates *learning* or the transmission of knowledge from incumbent exporters to potential entrants about specific aspects of the export markets, including information related to tastes, marketing practices, packaging, distribution channels, reputation of buyers etc. The downside of these mechanisms is the tougher competition on the (common) export market as well as potential congestion effects on input markets that may arise in denser areas (see for example Kang 2011).

Using transaction-level data for Peruvian exporting manufacturers this paper studies the importance of export spillovers from the agglomeration of exporters in a given region, on the extensive margin (the probability that an firm starts selling to a new export market) and on the survival dimension of the intensive margin of trade(the probability a newly created trade flow survives its first years of existence). Unfortunately empirically distinguishing the mechanisms described above would require highly detailed data which is rarely available. For this reason, this paper follows the empirical literature and focuses on studying the overall or net effects of these spatial externalities. These effects can come from a variety of sources and in principle can have both negatively and positive effects on entry and survival of export relationships, and therefore their net effects are an empirical question. Thus, the objective of this paper is to first test whether there is evidence of export spillovers affecting the entry and survival of new firm-level trade flows, and to explore the nature of these effects.

To motivate this empirical exercise I construct a simple extension of the model of heterogeneous exporters with idiosyncratic demand shocks developed by Cherkashin et al. (2015) to allow for uncertainty on fixed costs of exporting to a given market as in Freund and Pierola (2010) and Segura-Cayuela and Vilarrubia (2008), and to allow for a role for the geographic concentration of exporters in shaping this uncertainty. In the model the presence of other exporters in the same region make the access to foreign markets easier, and also can have positive effects on the survival of newly created trade flows. The model provides testable implications that fit well the patterns of creation and destruction of firm-level export flows observed in the data.

I find strong evidence that is suggestive of positive (net) local spillover effects from the proximity to other relevant exporters on both the probability that an existing exporter adds additional product-destination combinations to her portfolio, as well as on the survival prospects of these new export flows. These export spillovers appear to be both product and destination specific, and truly *local* in nature. Results are robust to the inclusion of several variables that control for other potential determinants of the entry and survival of new export flows, as well as to alternative specifications, and are both economically and statistically significant. The increment in the probability of exporting a new product-country destination from having an extra neighbor that produces and trades the same product to the same destination is 0.42 percentage points, all else equal. Although small, its magnitude gains relevance when compared this with the average unconditional probability of entering any product-destination in the sample (0.25 percent). In the case of survival, an additional peer on the vicinity selling the same product to the same country is associated with an increase of 1.18 percentage points on the probability of survival beyond the first year of a newly created export relationship (which on average survive with an unconditional probability of 45 percent).

This paper's contributions are threefold. First, I investigate the existence and nature of local export spillovers in the context of a developing resource-dependent country in which geographical and historical factors have determined a heavy concentration of economy production around a single area, the capital city. The city of Lima hosts more than 80% of manufacturers exporting in a given year while the rest are widely scattered in other urban areas across the Peruvian territory. This context contrasts with the less uneven spatial distribution of firms in the countries (mostly developed) in which export spillovers have been studied.[1] The fact that I find evidence suggesting the presence of export spillover effects, even when considering exporters located outside the capital city, reinforces the idea that these effects are indeed an important mechanism and therefore deserve more attention by both researchers and policymakers.

Second, I develop a simple partial-equilibrium extension of a model with heterogeneous exporters showing how local export spillovers can be introduced through trade costs and how

[1] Export spillovers have been studied in Spain (Moral-Benito 2013), France (Koenig 2009, Koenig et al. 2010), Belgium (Dumont et al. 2010), Russia (Cassey and Schmeiser 2012) and China (Fernandes and Tang 2014). In virtually all these countries, exporters tend to be concentrated in more than one area such as cities near important borders, or near important ports. In fact, with the sole exception of Koenig et al. (2010) all these studies focused on destination-specific spillovers only (i.e. aggregating away the product dimension of the data) precisely because it was evident ex-ante that one important reason for the agglomeration of exporting firms in different regions was proximity to the countries they served (for example, close to borders or along the coastal areas.). This feature is not present in the Peruvian data.

the empirical facts observed in the data can be explained using simple comparative statics. In the model firms are subject to uncertainty about the fixed costs of exporting, and this results in exporters with low productivity or with low demand shocks that decide to enter a given market, only to later discover that the return from this market is not high enough to survive longer than the first periods.

Finally, from a policy point of view the conclusions that can be drawn from this paper are relevant for the design of industrial policies and export promotion policies. In the first case, the conclusions from this paper can be relevant to answer questions about the effectiveness of cluster formation policies (e.g. the creation of industrial parks). In the case of export promotion policies, results presented here suggest giving more weight to the promotion of geographic clustering of exporters and to the creation of formal networks of firms to facilitate information sharing.

Relation to the literature. This paper related is to two strands of the international trade literature. First, this paper relates to the strand that studies export spillovers from the proximity to other exporters (or multinational firms) on the entry decision of domestic producers (e.g. Aitken et al. 1997, Clerides et al. 1998 Bernard and Jensen 2004). More recently a new group of papers have reexamine the issue using transaction-level data for different countries as well as finer definitions of export relationships (see Koenig 2009 and Koenig et al. 2010 for France, Dumont et al. 2010 for Belgium, Choquette and Meinen 2014 for Denmark.

Second, this paper also contributes to the literature that focus on the determinants of the survival of new trade flows (see Besedes and Prusa 2006; Gorg et al. 2008; Molina and Fugazza 2009; Esteve-Perez et al. 2011, and Albornoz et al. 2013). In particular, this paper adds to the new literature that focus on the role of export spillovers on the survival exports initiated by Cadot et al. (2011) and Fernandes and Tang (2014). Since the seminal paper by Besedes and Prusa (2006), the survival or duration of export relationships has received increasing attention as an additional dimension of export performance. Export survival is a key aspect of the internalization of domestic producers, as the success of this process must take into account the ability to keep selling to foreign markets over time (Molina and Fugazza 2009). Furthermore, Eaton et al. (2007), Albornoz-Crespo et al. (2010), Buono et al. (2008), among others have shown empirically that once an export relationship outlives its first years of existence, it exhibits substantial growth rates contributing greatly to the intensive margin and hence to the country's overall export growth. Also, as pointed by Manova and Zhang (2012), a good identification of the determinants of firms export success is key for the understanding of the patterns of international trade across countries, the welfare consequences of globalization, and the design of export-promoting policies that promote trade and ultimately growth in developing countries.

The rest of the paper is organized as follows. The next section presents the theoretical framework that is used to motive the empirical exercise. The empirical analysis is presented in section 3, starting with a description of the data set, a thorough description of the main specifications and some of the limitations faced. Empirical results are presented in section 4, followed by a discussion of the nature of the spillover effects found and a series of robustness checks. Finally, section 6 concludes.

2 Theoretical Framework

This section provides a simple extension of the standard model of heterogeneous exporters aimed at guiding the empirical analysis. The setting introduces two additional features to the multi-country version of the standard model of heterogeneous exporters: (i) firm and destination-specific demand shocks, a feature in recent models (see Eaton et al. 2009, Cherkashin et al. 2015, Crozet et al. 2012 among others); and (ii) the possibility that the proximity to other exporters in a given local area affects on the performance of new trade flows. The model is in partial equilibrium and takes the location of exporters as exogenously given. These features modify the productivity sorting mechanism characteristic to the Melitz model and hence provides a simplified way to explain the high rate of death in the initial years of life of new export relationships.

In this setting, the agglomeration of exporters in the same region influences both the sunk costs associated with drawing a demand shock and the fixed costs of exporting to a given export market faced by all firms located in a district. In the first case, these sunk costs are assumed to be a non-increasing function of the measure of exporters selling to that particular market in the region. In the case of fixed costs, these are are assumed imperfectly observed by new exporters as in Freund and Pierola (2010) and Segura-Cayuela and Vilarrubia (2008), and that their true value will only be revealed after the firm effectively exports for the first time. The probability of getting a good or bad realization is however linked to the prevalence of export operations in the district.[2]

2.1 Basic set-up

Consider a set of firms producing differentiated varieties indexed by ω in a given country. Preferences of the representative consumer in country j are given by $U_j = H_j^{1-\beta} C_j^{\beta}$: where H_j denotes consumption of an homogeneous good, while C_j is a composite of differentiated varieties defined as:

$$C_j = \left(\int_{\omega \in \Omega_j} \left[a_j\left(\omega\right) q_j\left(\omega\right) \right]^{(\sigma_j - 1)/\sigma_j} d\omega \right)^{\sigma_j/(\sigma_j - 1)} \tag{1}$$

where Ω_j is the set of differentiated varieties available in country j, σ_j is the elasticity of substitution between varieties in country j, and the $a_j(\omega)$ are variety- and country-specific demand shocks. Firm-destination demand shocks allow the model to accommodate the fact that two firms with the same level of productivity may differ both in the choice of countries they decide to export to and in the amount exported to the same country.[3]

Solving the representative's consumer utility maximization problem the demand for variety ω is $q_j\left(\omega\right) = \alpha_j p_j(\omega)^{-\sigma_j} E_j / P_j^{1-\sigma_j}$ where $P_j = \left(\int_{\omega \in \Omega_j} \left[p_j\left(\omega\right) / a_j\left(\omega\right) \right]^{1-\sigma_j} d\omega \right)^{1/(1-\sigma_j)}$ is the ideal aggregate price index associated with (1), $\alpha_j \equiv a_j(\omega)^{\sigma_j - 1}$, and $E_j = \beta Y_j$ is the portion of aggregate income Y_j spent on differentiated varieties.

The model adopts the timing of events of the production decisions used by Cherkashin et al. (2015), which explicitly separates the entry into the industry decision from the entry into an export market decision. First, entrepreneurs must pay a sunk cost to draw a productivity shock φ from the cumulative distribution function $G(\varphi)$ in order to enter the industry; firms will enter only if they draw a productivity shock larger than φ^* (industry entry decision).[4] Next in order to enter a foreign market j, firms must pay an sunk cost K_j

[2] This is consistent with the model developed by Albornoz-Crespo et al. (2010): entry is the only way firms can uncover information a given destination and then they can revise their profitability expectations for future entry and exit decisions.

[3] There are a number of possible interpretations for $a_j(\omega)$. In addition to cross-country variation in the tastes for variety ω, it could also represent a firms network of connections with purchasers in each market. Without demand shocks, the model would predict a perfect hierarchy of destinations according to which all exporters that serve a relatively though market (a remote or relatively small destination) should also serve all easier countries.

[4] See the appendix in Cherkashin et al. (2015) for a derivation of φ^*

to draw a demand shock α_j from the cumulative distribution $H(\alpha)$ (market entry decision). Finally given φ and α_j, the firm decides to effectively produce for exporting to destination j if and only if its demand shock is high enough (production decision).

The homogeneous good is produced everywhere under perfect competition and with a constant returns to scale technology. This good is freely trade across countries and will serve as the numeraire. The differentiated-good industry is instead monopolistically competitive. There are no economies of scope and therefore each firm produces a single variety and can be indexed by φ. Firms are heterogeneous both in terms of their productivity draws φ and, later, in terms of the country-specific demand shocks they receive. A firm with productivity φ incurs a per-unit cost of production equal to $1/\varphi$.

Trade is costly: in order to export to destination country j a firm must pay a variable trade cost, which takes the usual destination-specific iceberg cost form $\tau_j \geq 1$; and a destination-specific per-period fixed cost F_j, which can be interpreted as an overhead costs of exporting to j. Profit maximization implies the well-known result that prices will be equal to a constant-markup over marginal costs $p_j(\varphi) = (\sigma_j/(\sigma_j - 1))\tau_j/\varphi$.[5]. Using this pricing condition and the downward sloping demand curve it faces, a firm's profits from exporting to j are:

$$\Pi_j(\varphi, \alpha_j) = \alpha_j \varphi^{\sigma_j - 1} \Lambda_j - F_j$$

where $\Lambda_j \equiv \sigma_j^{-\sigma_j} (\sigma_j - 1)^{\sigma_j - 1} \tau_j^{1-\sigma_j} P_j^{\sigma_j - 1} E_j$ summarizes country j's market potential.

Without uncertainty, firms will start exporting to destination j as long as the present value of profits, discounted at a rate δ is non-negative: $\sum_t \delta^t \Pi_j(\varphi, \alpha_j) - K_j = (1/(1-\delta)) \Pi_j(\varphi, \alpha_j) - K_j \geq 0$. If a firm decides to stop exporting to j, then its option value is normalized to 0. Then for any firm with productivity $\varphi \geq \varphi^*$, the minimal demand shock to sell to destination j is defined by $\Pi_j(\varphi, \alpha_j^*(\varphi)) = (1-\delta)K_j$:

$$\alpha_j^*(\varphi) = \frac{(1-\delta) K_j + F_j}{\varphi^{\sigma_j - 1} \Lambda_j} \tag{2}$$

That is, the expression for the demand shock cutoff $\alpha_j^*(\varphi)$ is common to all exporters but its final value is firm-specific. Note that for highly productive firms, the cutoff $\alpha_j^*(\varphi)$ is low, implying that those firms will be more likely to start exporting to j. In turn, low

[5] Note that the price does not depend on the magnitude of the demand shock α_j

productivity firms require a high realization of the demand shock to start exporting to a given destination which make them less likely to serve more export markets. This feature is consistent with (i) more productive firms exporting to more and tougher destinations, (ii) some relatively less productive (and therefore smaller) exporters selling to more distant and tougher destinations, and (iii) some large exporters not selling to closer and larger (easier) markets. All these patterns are well known in the empirical literature (see for example Eaton et al. 2007, Amador and Opromolla 2008 and Lawless and Whelan 2008).

2.2 Local export spillovers

To introduce spatial externalities the set-up described above is extended in two ways. First, assume now that firms are spatially distributed in a finite number of districts indexed by ℓ, and let $z_{\ell j}$ be a measure of the prevalence of other exporters shipping their products to market j operating in district ℓ. Second, both the sunk costs K_j that potential entrants to market j must pay to draw their idiosyncratic demand shock α_j and the per-period fixed costs F_j faced by entrants will be affected by the concentration of other firms exporting to j in district ℓ.

In the case of sunk costs, these will now be allowed to vary across regions by assuming that they are a deterministic function of $z_{\ell j}$. Assumption 1 states differentiability and regularity conditions for $K_{\ell j} \equiv K(z_{\ell j})$:

Assumption 1. *The sunk cost* $K(\cdot)$ *is decreasing and differentiable with* $K' \leq 0$ *and* $K'' \leq 0$ *and satisfies* $\lim_{z\to 0} K_j(0) = \bar{K}_j$ *and* $\lim_{z\to +\infty} K(z) = 0$.

This assumption captures the notion that the agglomeration of exporters selling to j in ℓ will make it easier for potential entrants to effectively start exporting to the same export market j due to cost-sharing mechanisms or by facilitating the transmission of information about demand conditions.

On the other hand, as in Freund and Pierola (2010) and Segura-Cayuela and Vilarrubia (2008) entrants face uncertainty on the fixed costs of exporting to country j; however unlike these authors, the model here assumes that the proximity to other exporters play a role in shaping this uncertainty. Specifically, with probability $\rho_{\ell j} \equiv \rho(z_{\ell j})$ an entrant to export market j gets a low fixed-cost realization $F_j = F_j^L$ and with probability $1 - \rho_{\ell j}$ the firm that enters j will face high fixed costs $F_j = F_j^H > F_j^L$. Assumption 2 establishes some regularity conditions for the region- and destination-specific probability measure $\rho_{\ell j}$:

Assumption 2. *The probability measure $\rho(\cdot)$ is increasing and differentiable with $\rho' \geq 0$ and, $\rho'' \leq 0$ and satisfies the following regularity conditions: $\rho(z) \in [0,1]\ \forall z$, $\lim_{z\to 0}\rho(z) = 0$ and $\lim_{z\to +\infty}\rho(z) = 1$.*

Uncertainty is only revealed to the entrant after it effectively exports to market j; after that F^j will remain at that value forever.[6]

To see how the agglomeration of exporters and the uncertainty surrounding fixed costs affect entry and exit in region ℓ, we can analyze the first two periods of the model backwards. All results will be conditional on entry into the industry, i.e. for all firms for which $\varphi \geq \varphi^*$. In period 2, after the cost shock has been revealed, a firm located in ℓ must decide whether to continue exporting to j and making $\Pi(\alpha_j, \varphi)$ or exit which gives it a return of 0. The minimal demand shock required for a firm with productivity φ and cost-shock k to consider continue exporting to j is:

$$\frac{1}{1-\delta}\Pi_j(\varphi, \alpha_j^k) = 0 \ \Rightarrow \ \alpha_j^k(\varphi) = \frac{F_j^k}{\varphi^{\sigma_j - 1}\Lambda_j}, \quad k \in \{H, L\} \tag{3}$$

where $\alpha_j^H(\varphi) > \alpha_j^L(\varphi)$. Note that these ex-post cutoffs are completely independent of $z_{\ell j}$.[7]

In the first period, before uncertainty is resolved and after drawing their demand shock realization α_j firms must decide whether or not to enter country j by looking at the value of entry. As in Freund and Pierola (2010), there will be three types of firms. First firms with low demand shocks will find it optimal not to start exporting to j even if they get a low fixed costs realization. The value of not starting to export to country j is normalized to zero. Second, firms with intermediate levels of demand shocks will attempt entry and will choose to stay if and only if they receive a low realization of the fixed cost. The value

[6] Firms may decide to experiment certain export markets only to discover the exact nature of these costs, and withdraw if they are unexpectedly high. For the case of the first period, assume that exporters sign a contract to sell a given amount of the variety they produce which they cannot renege even if they realize that they will make negative profits.

[7] Positive per-period profits are a necessary but not a sufficient condition to ensure survival, so these cutoffs are only referential (because the sunk cost K_j must be taken into account.)

function for this type of firms is:

$$V_j^{E/X}(\alpha_j,\varphi,z_{\ell j}) = (1-\rho_{\ell j})\left[\alpha_j\varphi^{\sigma_j-1}\Lambda_j - F_j^H + \frac{\delta}{1-\delta}(0)\right] + \rho_{\ell j}\left[\frac{1}{1-\delta}\left(\alpha_j\varphi^{\sigma_j-1}\Lambda_j - F_j^L\right)\right] - K_{\ell j} \quad (4)$$

where the second term inside the first bracket highlights the firm's exit decision after exporting for one period (thus getting a zero payoff).

Finally, firms with a very high demand shock realizations will always decide to export to j, regardless of their fixed cost realization. The value function for this type of firms is:

$$V_j^{Always}(\alpha_j,\varphi,z_{\ell j}) = (1-\rho_{\ell j})\left[\frac{1}{1-\delta}\left(\alpha_j\varphi^{\sigma_j-1}\Lambda_j - F_j^H\right)\right] + \rho_{\ell j}\left[\frac{1}{1-\delta}\left(\alpha_j\varphi^{\sigma_j-1}\Lambda_j - F_j^L\right)\right] - K_{\ell j} \quad (5)$$

Equalizing the value function (4) with the value of not starting to export to country j (zero) gives the minimal demand shock for a firm with productivity φ to start exporting to country j:

$$\alpha_j^{**}(\varphi,z_{\ell j}) = \frac{(1-\delta)K_{\ell j} + (1-\delta)(1-\rho_{\ell j})F_j^H + \rho_{\ell j}F_j^L}{(1-\delta+\delta\rho_{\ell j})\varphi^{\sigma_j-1}\Lambda_j} \quad (6)$$

Similarly, the marginal exporter that is indifferent between staying forever in j and entering and ceasing exporting after the first year has demand shock equal to the ex-post cut-off α_j^H defined in (3).[8]

The relationships between these cutoffs for a firm with productivity $\varphi > \varphi*$ are illustrated in figure 1. All firms with $\alpha_j \geq \alpha_j^{**}$ will initially enter export market j. The two lines represent the value of entering to export market j and exiting if the fixed cost realization is high (5), and the value of continuing exporting regardless of the cost shock (4).

Because $F_j^H > F_j^L$ for all j it is easy to show that the ex-ante cutoff (6) will always be to the right of $\alpha_j^{*L}(\varphi)$ the demand shock cutoff under no uncertainty in fixed costs defined in (2), with $F_j = F_j^L$.[9] This result, illustrated in figure 1, has two implications: (i) firms

[8] This cutoff results from equalizing $V_j^{E/X}(\alpha_j^H,\varphi,z_{\ell j}) = V_j^{Always}(\alpha_j^H,\varphi,z_{\ell j})$

[9] It can be shown that this hypothetical cutoff is between $\alpha_j^L(\varphi)$ and $\alpha_j^{**}(\varphi)$ as illustrated in figure 1.

that start exporting to j with a low fixed cost realization will never exit, and (ii) there is a non-zero measure of firms that missed the opportunity to start exporting to a new export market profitably due to the uncertainty on fixed costs. To see this last implication, note that firms that drew demand shocks between $\alpha_j^{*L}(\varphi)$ and $\alpha_j^{**}(\varphi)$ and received a low cost realization would have started exporting to j profitably but chose not to do so due to the uncertainty in fixed costs.

On the other hand, the relationship between $\alpha_j^H(\varphi)$ and $\alpha_j^{**}(\varphi)$ will depend on the relationship between ρ, δ, K_j and the gap between F_j^H and F_j^L. To focus on the interesting case in which there is a clear hierarchy of cutoffs that generates exit, i.e. $\alpha_j^L(\varphi) < \alpha_j^{**}(\varphi) < \alpha_j^H(\varphi)$ in what follows I will assume:

Assumption 3. *In all regions the measure z is sufficiently above from 0 and the relative value of the difference between F_j^H and F_j^L and $1-\delta$ is large enough to guarantee the satisfaction of the strict inequality $\rho_{\ell j}\left(F_j^H - F_j^L\right) > (1-\delta)\, K_{\ell j}$.*[10]

Under assumption 3 those firms that get a high fixed-cost realization and demand shocks that lie between $\alpha_j^{**}(\varphi)$ and $\alpha_j^H(\varphi)$ will decide to exit after the first year of exporting to j.

Finally, note that the fact that all cutoffs move to the left when φ increases means that more productive firms will be more likely to export to more markets and at the same time they will be more likely to survive.

Effects of an increase in $z_{\ell j}$

In order to assess the effect of local export spillovers in the model, we can use comparative statics by looking at how the ex-ante cutoff (6) changes as $z_{\ell j}$ increases. The assumed differentiability of $\rho(\cdot)$ and $K_j(\cdot)$ simplifies the analysis by allowing us to take derivatives. Appendix A shows that an increase in $z_{\ell j}$ results in a reduction in the ex-ante cutoff, a reduction in the sunk costs of drawing a demand shock, and hence an increase in average entry.

Proposition 1. *Firms located in a region with a high prevalence of other exporters selling to j will be more likely to start exporting to j.*

[10] Note that when z is very high, then K approaches zero and ρ approaches one. Hence, the strict inequality is easily satisfied. In contrast, when z approaches 0, the ratio $\bar{K}/\rho$ goes fast to infinity and hence α_j^H will be to the left of the ex-ante cutoff $\alpha_{\ell j,t}^{**}$

Proposition 1 relies on two elements of the model. First, the sunk cost of starting to export to market j is assumed to be a decreasing function of $z_{\ell j}$ which means that the higher the number of nearby exporters that sell to j, the cheaper it becomes to draw a demand shock for potential entrants. Second, the ex-ante cutoff (6) is also a decreasing function of $z_{\ell j}$. This last result implies in turn that as the concentration of exporters selling to j in region ℓ increases, the ex-ante cutoff approaches α_j^L the ex-post cutoff for firms with low fixed costs (which is independent of $z_{\ell j}$). This means that the measure of firms that oversee the opportunity to profitably export to j due to the uncertainty will be lower, and hence entry should be higher on average. The additional firms that enter due to this reduction in $\alpha_{\ell j}^*$ will follow the entry-exit strategy.

Figure 2 shows this by illustrating how the curves $V_{\ell j}^{Always}$ and $V_{\ell j}^{E/X}$ change with an increase in z. The increase in z reduces the absolute value of the intercept of $V_{\ell j}^{Always}$ without affecting its slope, so that it registers an inward parallel shift; while it makes $V_{\ell j}^{E/X}$ steeper and changes the intercept of $V_{\ell j}^{E/X}$ so that both new curves intersect again on α_j^H.

Let $\xi_{\ell j}^H$ denote the measure of exporting firms that draw a demand shock between $\alpha_{\ell j}^*$ and α_j^H. A corollary of 1 proved in appendix A is:

Proposition 2. *The proportion of firms that engage in the entry-exit strategy is increasing in $z_{\ell j}$.*

The overall survival rate for firms that start exporting to market j from region ℓ can then be written as:

$$\Pr(\text{Survival}_{\ell j}\left(\varphi, \alpha, z_{\ell j} \left| \varphi \geq \varphi_j^* \right.\right) = (1 - \rho_{\ell j}) \cdot \left(1 - \xi_{\ell j}^H\right) + \rho_{\ell j} \cdot 1$$

where the first term is the contribution to the overall survival rate of firms that drew a high fixed costs realization (which survive with probability $1 - \xi_{\ell j}^H$), and the second one is the contribution of firms that received a good draw (which always survive). Appendix A shows that,

Proposition 3. *An increase in $z_{\ell j}$ causes a reduction in the contribution of firms with high fixed costs to the overall survival rate, while at the same time it increases the contribution of firms with low fixed costs (because these firms survive with probability 1 and their proportion will now be higher). Which effect dominates will depend on the initial level of $z_{\ell j}$ and the* *curvature of the $\rho(\cdot)$ function: as long as the increase in the probability of receiving a good draw that is caused by an increase in $z_{\ell j}$ is large enough, which given the assumed curvature of $\rho(\cdot)$ occurs when $z_{\ell j}$ is low, then the overall survival rate will be higher.*

This proposition implies that a reduction in the overall survival rate is feasible specially when $z_{\ell j}$ is high, which is consistent with the notion of congestion effects arising in denser areas described in the introduction. Empirically, this implies the existence of an inverted-U shape of the relationship between z and the average survival of entrants in export market j. These two results follow from the regularity conditions assumed for $\rho(\cdot)$: for a low (high) z, the slope of $\rho(z)$ will be high (low) and therefore as z decreases (increases) it will be more (less) likely to observe a higher (lower) survival of firms into j.

Finally note that this theoretical framework does not yield any testable implication for the initial size of each trade flow. This is consistent with the fact that we are only allowing for changes in the fixed/sunk costs of exporting: it is a well known result in the models of heterogeneous exporters (Chaney 2008), a change in the fixed component of trade costs only affects the extensive margin while a change in variable trade costs affects both the deepening dimension of the intensive margin as defined by Besedes and Prusa (2007) and the extensive margin of exports.

3 Empirical analysis

Following the predictions of the model outlined in section 2, this section analyzes the effect of export spillovers on the probability that a firm enters a newly product-destination export market (extensive margin) and the probability that this new export relationship survives beyond its first years of existence (the survival component of the intensive margin). The base spillover measure that will be used in the regression analysis below is the number of exporters whose plants are located in the same district.

Due to the strong concentration of exporters in the capital city described in the introduction, the study of export spillovers requires using a more disaggregated definition of a trade flow. In particular, unlike most studies on this topic[11] that use only the country dimension of the data (i.e. trade flows are defined there as a firm-country combination), this paper uses a more disaggregated firm-product-country definition.[12]

In the absence of an observed shock that could provide exogenous variation to obtain identification of the local spillover effects, the empirical strategy consists of using an comprehensive set of fixed effects to absorb all sources of unobserved variation that are not specific to a particular aspect of the theoretical framework outlined in section 2. In particular, firm-year fixed effects are used to control for unobserved firm-specific time-variant factors such as firm's TFP, age effects, capacity constraints and other idiosyncratic shocks but also for time-invariant factors such as the firm's manager ability to design a successful internationalization strategy for the firm. One advantage of using firm-year fixed effects is that their implementation does not wipe out the potentially rich information provided by the firm's unchosen potential choices and their relationship with the firm's local environment. The time-dimension of firm-year fixed effects also allows to account for more general factors such as exchange rate shocks and the overall macro environment. On the other hand, the inclusion of product fixed effects allows to control for time-invariant product characteristics such as the country's comparative advantage in the production and exporting of a particular product which is unlikely to change within the sample of years considered. Finally, region fixed effects allow to control for time invariant city characteristics that attract firms.[13]

The econometric analysis for both outcomes is based on a linear probability model. Despite its well-known caveats this is the most convenient choice to model entry and exit relative to non-linear specifications like logit or probit, mainly because it allows the inclusion of high-dimensional fixed effects (see Cameron and Trivedi 2010) a central component of the present analysis.

[11] See for example Moral-Benito (2013) who study the case for the Spanish exporters; Koenig (2009) and Koenig et al. (2010) who study the French case; Dumont et al. (2010) for the Belgian case; Cassey and Schmeiser (2012) for the Russian case, and Fernandes and Tang (2014) for the case of Chinese manufacturers. In all these cases it was evident ex-ante that one important reason for the agglomeration of exporting firms in different regions was proximity to the countries they served (for example, close to borders or along the coastal areas). This feature is not present in the Peruvian data.

[12] Koenig et al. 2010 focused on both destination-specific and product-specific spillovers using a firm-product-country definition of export relationship.

[13] More productive firms self-select and better able to survive in bigger cities due to competition effects that are not related to any kind of agglomeration gains (Melitz and Ottaviano 2008)

3.1 Data

The main data set employed in this paper is based on Peruvian transaction-level customs data collected by the Peruvian Customs office SUNAT-Aduanas (*Superintendencia Nacional de Administracion Tributaria*) that cover the universe of Peruvian exporters between 1996-2012.[14] The unit of observation is the export declaration form which includes the date of the transaction, the products exported, the destination country and the value of the sale among other details of each transaction. Firms are identified by their local tax identification number and products are classified according to the 10-digit Nandina product classification common to all members of the Andean Community the first six digits of which coincide with the Harmonized System.[15]

Data on firm-characteristics also came from SUNAT and includes each exporter's main business activity and the number of workers declared for social security purposes for each year between 2001 and 2011. Crucially for this paper, I had access to the exact location and type of *all* establishments associated with each firm (e.g. headquarters, plants, commercial shops, warehouses and administrative offices). This data is static in the sense that it reflects the most up-to-date information. However since the empirical setting focuses on single-plant manufacturers that registered exports in recent years I assume that the information on location obtained in 2012 applies to all years under analysis.[16]

Additional sources include data on world imports at the product level from U.N. COMTRADE and data on GDP, exchange rates are from the World Development Indicators of the World Bank. Finally, data on bilateral distance, common language and common border between Peru and its trade partners come from *Centre d'Etudes Prospectives et d'Informations Internationales* (CEPII).[17]

The data set is aggregated to yearly flows at the 4-digit HS product level.[18] In order to

[14] Access to the database was kindly provided by Sociedad de Comercio Exterior del Peru, Comex-Peru.

[15] This data set was subject to an deep cleaning process to minimize inconsistencies and cases of measurement error. Appendix B.1 describes this process.

[16] Plant reallocations are possible but must occur with a very low probability, but if it occurs most likely will be within the same district. Since the spillover measure is based on the concentration of exporters in a given district, this is a lesser problem. However in some cases in which there appeared to be some inconsistencies, it was possible to consult historical records on a particular firm's location. The cases of mismatch were minimal.

[17] This data is available at http://www.cepii.org.

[18] The reason to work with a rather aggregate product definition relies entirely on computational limitations, as described in section 3.

minimize inconsistencies due to changes across versions of the Harmonized System product classification I applied the product concordance developed by Cebeci (2012) that corrects for changes in the HS nomenclature across time.

The empirical analysis focuses on single-plant manufacturing firms exporting consumption and intermediate products.[19] A firm is considered to be a manufacturing firm if its reported main business activity corresponds to any of the manufacturing sub-sector of the ISIC revision 3 industrial classification, and did not report wholesaling or retailing as a secondary activity.

Table 1 present some basic statistics for the years in the estimation sample. In 2006, there were 1,828 manufacturers that exported almost 500 product codes in 72 industries to 141 destination countries. Single-plant manufacturers firms represented around 90% of all manufacturing exporting firms in all years under analysis as shown in table 2. In regressions both right-hand side and left-hand side variables refer to single-plant manufacturers.[20]

Exporters' spatial distribution. The Peruvian territory is divided in three administrative subdivisions, starting from the most aggregate: regions (*regiones*), provinces and districts.[21] The average area of a district in the country is 700 squared kilometers, but it is considerably lower in urban areas (in the city of Lima the mean district has an area of only 58 square kilometers.) Even though I can observe the exact addresses of all establishments associated with all exporters, the lack of an efficient post-code system for local addresses makes it impossible to geocode them using standard techniques.[22] For this reason, the analysis will use the district as the most disaggregated area to study spillover effects.

The spatial distribution of single-plant manufacturers that registered exports in 2006 across districts is shown in figure 3. As a reference point to illustrate the unequal distribution of economic activity across the Peruvian national territory, this map also shows the location of all plants owned by firms that are listed at the National Registry of Manufactured firms constructed by the Ministry of Production, which in turn is based on the registry for the

[19] Capital goods are excluded as several export transactions were identified as corresponding to the occasional selling of assets by firms that are reallocating or were in the process of being liquidated.

[20] A similar strategy was used by Koenig et al. (2010) to study the effect of local spillovers on the decision to start exporting by domestic French producers.

[21] In the period of analysis, there were 24 regions, 194 provinces and around 1800 districts nationwide.

[22] It is however possible to geocode addresses located in main cities, particularly the capital city Lima. However the number of addresses for which their coordinates can be accurately retrieved (for example using apps that use Google Maps) is too small.

National Economic Census conducted in 2007. Circles represent different bins for the number of exporting single-plant manufacturers located in a given districts, adjusted for the area of the district. There are two patterns that are worth highlighting. First, as anticipated exporters are heavily concentrated in the Lima-Callao area in the center coastal area.[23] In a given year about 75 percent of all exporters were located in the Lima-Callao area, while goods shipped from the sea and air customs located in El Callao represented roughly 80 percent of the aggregate value. Second, there are certain regions outside Lima-Callao which exhibit some degree of agglomeration of exporters, specially in coastal regions in northern Peru. Finally, outside Lima, districts that host exporters tend to be widely dispersed and there are large areas were no exporters were located at all.[24]

The map also zooms over the Lima-Callao region to how plants are agglomerated in the districts that conform Peru's capital city. This figure shows that within this region, there are some districts that host a great number of single-plant exporters, specially in districts near the area in which the country's main port and airport are located (next to the peninsula that appears at the centre of the map). Across groups of products, the concentration of plants in the Lima-Callao region may vary as is shown in figure 4.

3.2 Entry into new export markets

An ideal setting to study the entry decision of firms into new export markets would include the observation of all firms at risk of entry as well as their set of potential choices. In any given year, the set of firms at risk of entry or potential entrants can be composed by incumbent exporters, by first-time exporters and by non-exporters. Unfortunately, observing non-exporters for the Peruvian case is not possible due to the lack of a census or a representative survey of manufacturing firms that could be linked to the customs data. Additionally defining the set of potential choices requires knowledge of he previous export behaviour of firms. For these reasons, the analysis will focus on the entry decision of incumbent exporters

[23] The Lima-Callao area consists of core districts of the Province of Lima and the special province of El Callao–a subset of districts within the Province of Lima that gained the statue of Province after Peru's independence from Spain. The country's main seaport and airport are both located in the El Callao.

[24] Historical and geographical factors explain this situation. On the one hand, since the time in which Peru was a colony of the Spanish empire the economic activity and urban areas have been heavily centralized around the capital city and near seaports along the coastal line. On the other hand, the geography of the Peruvian territory is particularly complicated as the Andes mountains cut across the country from northwest to the southeast, representing (even today) a severe constraint for the development of transport infrastructure and hence for the development of productive areas at the interior regions.

to new product-country combinations in a given year, that is manufacturers that became exporters at least the year previous, leaving aside those firms that export for the very first time are excluded from the analysis.[25]

To define the set of potential choices faced by entrants, a number of restrictions must be applied to the data in order to avoid the unnecessary (and uninformative) growth of the number of observations as well as to avert computational limitations for estimation. First, as described before a rather aggregate definition of a product is used, based on the first four-digits of the HS product codes.[26] Second, the analysis is restricted to the period 2002-2007 in order to focus on a period in which Peruvian exports of manufactured products exhibited strong positive growth rates. Third, to define the set of potential countries that can be served by a firm located in district ℓ, only countries ever served by firms located in the same district in any year of the full-sample of the export transaction data (between 1996 and 2012) are taken into account. Finally, to define a firm's set of potential products only 4-digit HS codes exported by other manufacturers located in the same region that belong to the set of 2-digit HS industry codes exported by firm f the year previous are taken into account.[27]

Applying these restrictions, the number of observations increases significatively from 25 thousand to around 13 million observations between 2002 and 2007. While most of these observations corresponds to zero-trade flows, they provide potentially valuable information regarding the effect of local export spillovers on exporters' entry choices after controlling for other key determinants of entry.

The dependent variable for entry is constructed as follows:

$$\text{Entry}_{f\ell pd,t} = \begin{cases} 1 \quad \text{if}\{\, V_{f\ell pd,t-l} = 0\}_{l\geq 1}\ , \quad \sum_{h,k} V_{f\ell hk,t-1} > 0, \quad V_{f\ell pdt} > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad \text{otherwise} \end{cases} \tag{7}$$

where $V_{f\ell pd,t}$ denotes the value of exports of product p to country d by firm f located in

[25] A similar strategy is used by Araujo et al. (2013) using Belgian customs data. In unreported regression, new exporters were included in the sample and estimation results did not change significatively.

[26] That is, for instance this considers that exporters selling 6-digit product code 610412 "Women's or girls' suits of cotton" can be influenced by exporters selling the same product code but also by firms exporting product 610433 "Women's or girls' Jackets and blazers or synthetic fibres", both of which belong to the same 4-digit code 6104 "Women's or girls' suits, ensembles, jackets, blazers, (...) knitted or crocheted".

[27] Paravisini et al. 2011 defines potential product-destination choices, also for Peruvian exporters, using similar restrictions to the data in order to study the effects of bank credit on the extensive margin of exports during the recent financial crisis.

district ℓ at time t. The indicator function $Entry_{f\ell pd,t}$ equals 1 if firm f starts exporting product p to country d for the first time in year t. Since the Customs data goes up to 1996 and care was taken to guarantee the comparability of products across years, I can be confident that an observed entry is really the first time that an exporter is selling a product-country combination.[28]

The baseline estimation specification is then:

$$\Pr\left(\text{Entry}_{f\ell pd,t} = 1 \middle| \mathbf{X}\right) = \lambda_{ft} + \lambda_p + \lambda_\ell + \theta Spill'_{\ell pd,t} + \mathbf{Z}'_{pd,t}\beta_1 + \mathbf{W}'_{fpd,t}\beta_2 + \upsilon_{f\ell pd,t} \tag{8}$$

where λ_{ft}, λ_p and λ_ℓ denote firm-year, product and district fixed-effects. The matrix of country-product-year characteristics $\mathbf{Z}'_{pd,t}$ includes the aggregate volume of imports of good p by country d at time $t-1$ from *all* trade partners as a measure of product-specific worldwide demand, and the geodesic bilateral distance between Peru and country d as well as dummies for common language and sharing a border as proxies for trade costs. On the other hand, the matrix of firm-product-country controls $\mathbf{W}'_{fpd,t}$ includes measures of experience of the exporter with the country (the number of other products exported to country d) as well as with the product (the number of countries to which product p is also exported by the firm). Finally, to account for the sequential expansion of export markets and for the fact that export experience facilitate entry in geographically or economically similar destinations, for each product-country combination I include the count of countries to which the firm exported product p that either share a border with product d, or use the same official language, or share a border. That is, these variables control for *extended* gravity forces as defined by Morales et al. (2015).

The parameter of interest here is θ which gives the marginal effect of an increase in the measure of the prevalence of export operations in region ℓ at $t-1$ on the probability that a firm decides to start exporting a new product-country combination.

3.3 Survival beyond the first year

As discussed above, the model does not predict any effect on the traditional intensive margin (the value of each trade flow). However, it does show that survival dimension of the intensive

[28] Table A1 illustrate the distribution of the dependent variable for entry in 2006 according to the number of neighbors of exporters in the sample.

margin can be influenced by these local export spillovers. The variable of interest for this purpose will be the probability that a newly created export relationship survives beyond its first year of existence.[29]

Let $\text{Survival}_{f\ell pd,t}$ be an indicator variable that takes the value of 1 if a firm f located in district ℓ starts to export p to country d at time t for the first time, and survives at least until $t+1$. The dependent variable is then defined as follows:

$$\text{Survival}_{f\ell pd,t} = \begin{cases} 1 & \text{if } \{V_{f\ell pd,t-l} = 0\}_{l\geq 1}, \quad V_{f\ell pd,t} > 0, \;\; V_{f\ell pd,t+1} > 0 \\ 0 & \text{if } \{V_{f\ell pd,t-l} = 0\}_{l\geq 1}, \quad V_{f\ell pd,t} > 0, \quad V_{f\ell pd,t+1} = 0 \end{cases} \tag{9}$$

That is, a new firm-product-country trade flow is considered a *success* if it lasts at least two years, and a *failure* otherwise.[30]

The specification for estimating the effects of local export spillovers on the probability that a new export relationship survives beyond its first year of existence is:

$$\begin{aligned} \Pr\left(\text{Survival}^k_{f\ell pd,t} = 1 \middle| \mathbf{X}, \text{Entry}_{f\ell pd,t} = 1\right) &= \eta_{ft} + \eta_p + \eta_\ell + \theta_2 Spill_{\ell pd,t} \\ &\quad + \mathbf{Z}'_{pd,t}\gamma_1 + \mathbf{W'}_{fpd,t}\gamma_2 + \varepsilon_{f\ell pd,t} \end{aligned} \tag{10}$$

where η_{ft}, η_p, and η_ℓ denote firm-year, product and district fixed-effects, respectively. Matrices $\mathbf{Z}_{pd,t}$ and $\mathbf{W}_{fpd,t}$ are defined as for the specification of entry: the same covariates that affect entry are likely to also drive the probability of survival. The only difference with respect to (8) is that here $\mathbf{W}$ includes the size of firm's f initial sale of product p to country d, which has been shown to be an important determinant for the probability of survival of trade flows(Eaton et al. 2012, Rauch and Watson 2003). That is, relationships that initiate with large initial values are less likely to cease after the first years relative to relationships that start with low values.

[29] The approach of modeling the survival past the first year is also used by Cadot et al. (2011) to study the survival of export relationships initiated by African exporters; by Araujo et al. (2013) to study the role of institutional strength in shaping firm export dynamics (although without incorporating the geographical dimension) and by Fernandes and Tang (2014) who analyze the role of geographical information spillovers among Chinese exporters.

[30] Table A2 illustrate the distribution of the dependent variable for survival in 2006 according to the number of neighbors of exporters in the sample.

3.4 Estimation issues

Before analyzing the estimation results it is worth discussing some well known estimation issues that are common to empirical applications on firm-level export dynamics. The positive correlation between alternative spillover measures and the probability of entry and the probability of survival beyond the first year of an new export relationship is suggestive of the presence of local export spillovers but it is not enough to claim causality. The identification of export spillovers on the probability of entry and survival of newly created export relationships requires an assessment of econometric issues related to reverse causality and simultaneity biases, omitted variable bias and sample selection bias.

First, reverse casuality implies that both entry and survival of a newly created relationship by firm f can affect the stock of exporters in a given district r. Similarly there can be simultaneity issues can affect the relationship between spillover measures and the survival of new export relationships as the former can be determined by the same factors as the probability of survival of a new export relationship. To make up for these potential problems I follow the usual practice in the empirical literature of lagging all right-hand side variables by one period.

A second set of issues are related to potential omitted variable bias. Even though the comprehensive set of fixed-effects described above which include firm-year, product and industry individual effects allows to partially offset this problem, there could still be factors that vary at other dimensions such as country-district and at the district-year level that could be affecting the survival of trade flows. In particular, district-year characteristics (e.g. changes in the local business environment) could be important factors that may not be properly controlled for with the benchmark specification. For this purpose I run robustness check with different sets of fixed-effects. As will be detailed in subsequent sections, all results are generally robust to these alternative specifications.

Finally, there can be concerns about selection bias affecting the estimation of the probability of survival of new export flows since the profitability of uninitiated flows is by definition unobserved. While the most appealing solution to deal with this problem would be apply an estimator for selection-bias in the spirit of Wooldridge (1995) which allows to control for unobserved heterogeneity in a panel as applied for example by Araujo et al. (2013), doing this with the unit of observation defined as the triad firm-product-destination would be computationally unfeasible. As a second-best approach, the robustness of the main estimates is assessed using different sub-samples of firms according to their size.[31]

4 Results

This sections presents the estimation results of the regression analysis outlined above. The focus is on the spillover measures which is a measure of the prevalence of export activities in the neighborhood of a given firm. If there local export spillover have non-negligible effects of the performance of new firm-product-country trade flows, then we can expect that such measures have a positive and significant effect on the probability of entry and of survival beyond their first years of existence of these new export relationships. Because the regressor of interest (the spillover measure) is at the district-market level, all tables include standard errors clustered at the district level (Moulton 1990; Gormley and Matsa 2013). All tables described below report the value of estimated parameters multiplied by 100 for ease of interpretation.

4.1 Effects on the entry decision to new export markets

I estimate some variations of equation (8) and collect the results in table 3. Column (1) includes only the most specific spillover measure, the number of exporters selling the same product to the same country the previous period, and the set of fixed effects. The value of the estimated parameter is positive and highly significant. The next two columns progressively add relevant controls, first for country-product specific demand and geographic frictions that are well known determinants of trade flows, and later measures of the experience as a exporter specific to the firm initiating a new export relationship. Column (2) adds country d's specific demand for imports of product p from all countries in the world (except Peru) as a measure of demand; the geodesic distance between country d and Peru; and dummies for when the trade partner shares a border or the same official language (in this case, Spanish). The effect of demand is positive and statistically significant, suggesting that everything else

[31] A third possibility could be to resort to the strategy used by Araujo et al. 2013. These authors employ a similar empirical strategy for equations of entry and survival, which includes firm-year fixed effects and show that a simple parametric assumption regarding the correlation between the firm-year fixed effects in each equation suffices to obtain consistent estimates for the coefficients of both equations by running them separately and applying the corresponding within transformation to remove these effects.

equal a higher demand for imports of product p increases the probability that an incumbent exporter starts selling the product to that country. Similarly, the effect of geographic distance is negative which suggests that entry into farther countries is less likely. The dummies for common border and language are both positive, but only the first is statistically significant. Column (3) adds the number of 2-digit HS codes exported to the same country, as a proxy for the experience of the firm with the country; the number of other countries to which product p is also exported, as a measure of firm's experience with the product; and finally the sum of all countries which share similar characteristics than country d (including same language, common border, same colony origin, same continent) to control for extended gravity forces. The effect of adding these controls on the estimated value of the spillover measure is a small reduction in its value, leaving it positive and statistically significant at 1%.

Less specific measure of spillovers are tested in column (4) and (5): first the number of manufacturers located within the same district exporting product p to any destination, and then the number of manufacturers exporting any product to country d also located in the same district. The value of the parameters are also positive, statistically significant but considerably lower in magnitude. These results confirm the finding of Koenig et al. (2010) in the sense that the export spillovers effects are both product *and* country specific.

Finally, column (6) retakes the most specific spillover measure and adds control for the number of other firms selling the same product to other countries, as well as the number of other exporters selling other products to the same country under analysis. The estimated coefficient of these last measures are tiny but not statistically significant and the magnitude of the coefficient is not far from the specification in which only the most specific spillover measure is included. For this reason, column (3) will be the benchmark results for the rest of the analysis: an additional peer in the vicinity exporting the same product to the same destination increases the probability of entry for the average exporter by 0.42 percentage points. The estimated impact seems small in magnitude but is comparable to what was obtained by other authors looking at different countries. Also, this effect is large compared to the unconditional probability of starting to export a product-country destination in any year, which is 0.25 percent.

The nature of the effect of export spillovers on entry

Table 4 explores whether the identified export spillovers are heterogeneous across some firms characteristics and gives some hints about the nature of these effects. To ease comparison with the base results, column (1) replicates column (3) of table 3.

Heterogeneous effects. Column (2) of table 4 explores the heterogeneity of relationship between the spillover effects and the size of firms, by adding an interaction term of the spillover measure with the exporter's size as measured by its total exports.[32] The value of the coefficient of the interaction is small, negative and marginally statistically significant suggesting that the spillover effect is somewhat stronger for smaller firms.

Column (3) tests for differentiated export spillovers on *young* exporters i.e. firms that exported for the first time on the previous period. The coefficient of the interaction term is positive, statistically significant and more than 1.5 times larger than the base effect on "established" exporters, suggesting that the influence of the spatial externalities is stronger on firms that are less experienced as exporters. This is a fairly intuitive result: older and more experienced firms have well established relationships with their input suppliers and are less likely to benefit from new information about export markets, and so are less likely to benefit from having more neighbours than exporters with little recent experience.

Quadratic effects. As was stated in the introduction, the spillover effects that this and similar studies uncover are really the *net* effects from the interaction of both positive and negative factors that arise from the proximity to other manufacturers selling to the same export market. A positive coefficient for the spillover measure means that the average contribution of positive factors behind the spillover effect dominates that of negative factors. However, as shown in the theoretical framework, it could be the case that factors from the negative side begin to dominate once the agglomeration of relevant exporters is high enough. In other words, factors like congestion in the use of infrastructure services and competition for local inputs could represent a severe bottleneck for the performance of exporters when the number of exporters is above a certain level (see Fujita and Thisse 2013, p. 130).

To check whether the relationship between the probability of entry and spillover effects exhibit an inverted U-shape, column (4) adds the squared of the spillover measure. The estimated coefficient obtained is negative, highly statistically significant but very small in

[32] Similar results are obtained when using the number of workers as a measure of size.

magnitude, suggesting that while there is room for congestion effects in regions with a very high concentration of firms, the effect is very unlikely to dominate positive effects from other mechanisms.

Spatial decay. Columns (5) and (6) of table 4 test how geographically specific are the spillover effects. Because the most disaggregated geographical area considered above is the district level, I check whether the spillover effects are constrained to firms located within the same district only, or firms located in other districts within the same province can also have an effect. That is, this is checking whether the local export spillover effects are subject to spatial decay.[33]

To do this equation (8) is reestimated using first the spillover measure calculated at the province level, and then distinguishing between exporters located in the same district and exporters located in other districts, but within the same province. If the parameter of the spillover measure for firms located in other districts within the same province is positive and statistically significant, but lower in magnitude than the coefficient for the spillover measure at the district level, then this would be evidence of spillovers exhibiting spatial decay.

First, column (5) adds the spillover measure but at a the province level, resulting in an estimate that although statistically significant and positive is an order of magnitude smaller than the estimate obtained with the number of exporters located in the same district. Column (6) adds to the benchmark measure of spillovers, the number of firms selling the same product to the same destination located in other districts within the same province. The resulting estimates suggests that these spillovers are highly localized and indeed subject to spatial decay since the value of the coefficient decreases with distance from the initial firm. The effect is almost 4 times smaller than the benchmark result.

Time decay. In their analysis of export spillovers on the export decision of UK firms, Greenaway and Kneller (2008) argued that spillover effects could be subject to decay through time. That is, focusing on the information sharing component of the spillover effects, the information from the recent entry experience of neighbors (i.e. the year previous) may be more relevant for potential entrants than information from exporters that entered long before. This situation could signal a potential problem if the whole effect is driven by the presence of new exporters as the coefficient might be picking up a contemporaneous exogenous shocks

[33] Choquette and Meinen (2014) and Koenig et al. (2010) find clear evidence of spatial decay when analyzing the role of export spillovers on the export propensity of local Norwegian and French producers, respectively.

on trade policy or foreign demand, that are not well controlled for by fixed-effects.

To check whether the local spillover effects found before are subject to this issue, I disaggregate the source of spillovers between export relationships initiated the year previous and those initiated on later years. In principle, this requires the assumption that the information sharing component of the spillover effects found dominates.[34] Here *new* exporters at time τ are firms that started to export product p at time $\tau - 1$ and then continued exporting at time τ. That is, an neighbor is new if it has at most two years of experience with product-destination pd. Results shown in columns (7) of table 4 indicate that although the effect on entry is stronger from new exporters, the presence of more experience exporters is also important and the effect is both positive and statistically significant.

4.2 Effects on the probability of survival beyond the first year

Results from estimating the linear probability model for survival outlined in (10) are presented in table 5. Using a similar strategy as for the case of entry, I first run the regression with the most specific spillover measure (the number of manufacturers exporting the same product to the same country on $t-1$), and then add different controls in the adjacent columns. A regressor added here that was not present in the estimation for entry is the initial size of the export relationship which has been identified as a key determinant of its probability of survival of trade flows (**?**). Results indicate that just as was the case for entry, the value of the estimated spillover effect is reduced in magnitude when controlling first for demand and geographic frictions, and then for the exporter's past experience with similar product or similar countries. The effect, however, remains positive and highly significant.

The specificity of the spillover effect on survival is checked as for entry using more aggregate measures in columns (4) and (5). In both cases, the value of the coefficient is much lower in magnitude and only with the number of firms selling any product to the same country, the coefficient is statistically significant. When including both aggregate measures and the most specific measure, then neither of the aggregate measures of spillover effects are statistically significant. For this reason, as was the case for the regression of entry, column (3) will be taken as the benchmark results for the remaining analysis. This suggests that an additional

[34] Other components of the spillover effects could also vary with the experience of incumbent exporters. For example, it also could be the case that more experienced exporters have well established relationships with local input suppliers and therefore, they are more likely to create congestion effects on this dimension.

peer on the vicinity selling the same product to the same country represents an increase of 1.18 percentage points on the probability of survival for the average newly established export relationship.

The nature of the effect of spillovers on survival

Heterogeneity across firm characteristics and a preliminary assessment of the nature of these spillover effects is studies in table 6. As for entry, column (1) replicates column (3) of table 5 for ease of comparison with the other results.

Heterogeneous effects. Column (2) explores the heterogeneity of relationship between the spillover effects and the size of firms that start selling new product-market combinations, as measured by the total size of their export sales. Unlike the case of entry, the value of the coefficient of the interaction of this measure of size is positive and statistically significant, while the main spillover effect is reduced in magnitude and ceases to be statistically significant. This reversal of results with respect to what was obtained for entry is somewhat puzzling as it reads as if the spillover effects on survival are stronger for larger firms only. This however can be interpreted as picking up part of the effects of total factor productivity on the survival of new export relationships that are not well controlled for with firm-year fixed effects: larger, more productive, firms will tend to survive longer on average regardless of the environment where they develop their productive activities.[35]

Column (3) tests the differentiated effect of the spillover measures for young exporters i.e. firms that exported for the first time on the previous period. Although positive, the coefficient of the interaction term is no statistically significant suggesting that there is no clear additional effect on survival for young exporters relative to that for more experienced ones.

Quadratic effects. Congestion effects could mean difficulties in the timely delivery of goods, which buyers can perceived as a signal of quality, and therefore can produce early relationship breakups. Similarly, the limited availability of quality inputs due to congestions, can also be translated into a lower quality of products being traded and therefore can affect the survival prospects of new trade flows that involve these products.

[35] This also can be linked to one of the predictions of the model which says that for any productivity level, the higher the concentration of other exporters, the more likely a firm is to enter. However, survival will only be positively affected as long as $z_{\ell j}$ is relatively low.

To check whether this is something that appears in the data (i.e. an inverted U-shape of the relationship between the survival prospects of new export relationships and the spillover measure), column (4) includes quadratic terms for the spillover measures in the benchmark specification. As was the case for entry, for survival the quadratic terms is negative and statistically significant although low in magnitude, suggesting that negative mechanisms behind the spatial externalities from proximity may being to dominate positive ones, and thereby generate negative spillovers, only on districts that exhibit a very high concentration of other exporters selling the same product to the same destination.

Spatial decay. Columns 5 and 6 of table 6 explores whether the spillover effects on survival are subject to spatial decay. Essentially here evidence for spatial decay is not as strong as for the case of entry. The coefficient for the measure of firms selling the same product to the same district located in any of the districts of the same province where the manufacturer under analysis has its plant is positive and considerably lower in magnitude than the specific effect shown in column (1). However, when splitting this measure into the number of firms located in the same district and the number of firms located in other districts within the same province, the estimated coefficient for the first measure is positive and large in magnitude but not statistically significant, while the coefficient for the measure that refers to firms located farther from the firm under analysis is statistically significant. In other words, the spillover effect appears to decay with distance but the main effect is not longer statistically significant.

Time decay. Finally, columns (7) explores time decay of the spillover effects. Estimates suggest that the spillover effect on survival is dominated by more experienced neighboring firms as the coefficient for the spillover measure corresponding to the number of these firms is large, positive, highly statistically significant and close in value to the main effect in column (1), while the effect from less experienced exporters is lower in magnitude but no significant. This suggests that these effects are not subject to time decay and what is more, for survival the effect of the nearby presence of experienced exporters is much more important for the survival prospects of newly formed export relationships for the average exporter.

4.3 Robustness checks

I perform a series of robustness checks to the main results described above. These checks include using alternative measures of spillovers, changing the definition of survival, using different sub-samples, and using different fixed-effects specifications.

Alternative measures of local spillover. The first robustness check comes from using alternative measures of local spillover. First, I use the aggregate value of export sales by firms selling the same product to the same country, located in the same district. Results in table A4 and table A6 reproduce the same results shown in tables 3 and 5 mimic well the ones obtained using the number of firms selling the same 4-digit product: the estimated coefficients are always statistically significant and positive.

A concern about using the number of neighbors exporting the same product to the same market is that this measure does not take into account the location of exporters within the district. In an ideal setting, the exact geocoded location of firms (or plants) would be available and this would allow the researcher to compute a much more accurate measure of the agglomeration of firms (e.g. the average distance to neighbors).[36] Unfortunately, for the Peruvian case this data is not readily available, in particular for firms located outside the main capital city.[37]

For this reason, some authors have used instead a measure of the density of firms in a given district, i.e. the ratio of the number of exporters to the area of the district. For this reason, table A3 replicates the base results shown in table 3 but with the spillover measure adjusted for the area of the district. All previous results are confirmed, implying that the number of neighbors is a valid spillover measure.

Different subsamples. The first five columns of tables A7 and A8 reproduces the base estimation using different subsamples of firms according to their size, as measure by their number of workers.[38] The first two columns use the size of the median exporter (10 employ-

[36] This is done, for example, for a cross-section of Spanish exporters by Moral-Benito 2013.

[37] The data set of addresses of establishments associated with exporters used in this paper does include sub-zones which are not formally nor legally well defined including "sub-urban areas", "business parks" or "industrial zones", for roughly half of firms included in the sample under analysis. Even though in principle firms could be grouped using these areas as a fourth level of disaggregation within the district, the fact that some of these areas overlap between two districts, and that not all addresses include such identifiers limits their use.

[38] While this is a measure of formal employment by these firms, and not the overall size of their labor demand, the fact that exporters tend to be more formal somewhat ameliorates concerns about the validity of this measure as a proxy for firm size.

ees) as a cutoff to distinguish large and small firms. Results are positive, highly statistically significant and their magnitude is very close to each other. Relative to the estimated values in the benchmark results, for the case of entry the estimated values are higher, while for survival the value of the coefficients they are very close.

Columns (3) and (4) use the approximate 75 percentile (50 employees) to create to sub-samples of firms. Same conclusions as above apply here. The only difference is that for survival, while results are similar to the ones used the median as the cutoff, their statistical significance is somewhat reduced.

Concentration in the capital city. Another potential concern about the results presented above is that the high concentration of the country's economic activity around the Lima-Callao region (see subsection 3.1) could mean that the spillover effects found above are driven entirely by firms located in this region. For this purpose I re-estimate results that appear in the third column of tables 3 and 5 using both the subsample of single-establishment firms located in the Lima-Callao region, as well as the subsample of single-establishment firms located in districts outside this area.

Results are shown in columns (6) to (8) of tables A7 and A8. First, column (6) and (7) run the regressions for the sub-samples of firms locates in Lima and outside Lima, respectively, suggesting a much more pronounced effect of the spillover externalities on firms located outside the country's capital. Only the regression corresponding to entry of firms located outside Lima fails to be statistically significant. All patterns found in the basic specification and in previous robustness checks are repeated when analysing these sub-samples. Direct marginal effects are all positive, statistically significant and larger in absolute value than the ones obtained using the whole sample of firms. In particular, when using the subsample of firms located outside the Lima region, the magnitude of the estimated coefficients is considerably larger, and also externalities from the proximity to intermediaries are now considerably larger than the spillover from the proximity to other manufacturers.

Changing the definition of a surviving relationship. The definition of a surviving export relationship (as one that survives for at least 2 years) used in the basic results described above is somewhat arbitrary. The analysis of the survival probability of new export relationships in subsection 3.1 shows that the great majority of new export relationships are terminated on their first years of existence. However the shape of the survivor function really "stabilizes" after the first three or four years. Therefore, here I repeat the exercise using stricter definitions of a successful export relationship.

Results are shown in table A9. Compared to the benchmark estimation results in columns 3 of table 5, estimated coefficients using the above stricter dependent variables show similar patterns. In all cases, estimated coefficients have lower absolute value but remain significant at the 10% and with the same sign as for the benchmark case. This means that the spillover effects found above are not driven by the number of new export relationships that last for exactly two years.

Alternative fixed-effects specifications. Finally, I run alternative versions of equations (8) and (10) with different sets of fixed-effects. First, in order to take into account factors that affect both outcome variables and that may be industry and district specific, I include firm-year fixed effects along with industry-district fixed effects. The idea is to make sure that the effects found are not driven by that shocks that affects specific industries within a region. Second, similarly the specification could be the result of factors that are both country and district specific, I include firm-year fixed effects, product fixed effects and district-country fixed effects. Third, an even stricter version of fixed effects include firm individual effects to account for unobserved time-invariant characteristics, and industry-district-year fixed effects, which account for time-variant effects that vary both across districts and industries. Last, there can also be concerns that shocks are common for all firms exporting the same product from a given district, and so I include firm, product and district-year fixed effects.

Results for the probability of entry and survival are shown in tables A10 and A11, respectively. The first column of both tables repeat the baseline results obtained in column (3) of tables 3 and 5, respectively. In both cases, results are fairly robust to the different fixed effect specifications, remaining statistically significant and without major variations in absolute terms.

5 Concluding Remarks

This paper investigates the existence of local export spillovers on new export relationships initiated by Peruvian manufacturers, that is whether the proximity to other exporters exerts a non-negligible influence on the probability that a manufacturer starts exporting a product-country combination for the first time, as well as the probability of survival of this newly created trade flow. Results point towards a strong correlation between the concentration of exporters in a given district and both the probability of adding a new product-country combination to the firm's export portfolio, and the survival prospects of these new trade flows.

At first sight Peru may not appear as the best place to study these kind of effects. Unlike other countries studied by other authors in which the spatial distribution of firms is less uneven, in Peru the country's economic activity is heavily concentrated in a single area (the capital city). However even under these circumstances the evidence for local export spillovers on both margins of exports considered here is strong and holds to some extent even when looking at firms outside the main city.

The paper also explores the nature of these spillover effects. First, there is some evidence on the heterogeneous effects in terms of the size of the exporter and in terms of the experience of the firm at the time of exporting a new product-country for the first time. Second, these spillovers appear to be really local in the sense that proximity to firms within the same local area matters the most. Third, there is no clear indication that the effects are subject to time decay, in the sense that only the newest export relationships in a given area are the ones driving the spillover effects.

Results presented here can be potentially useful for the design of industrial and export promotion policies. In the first case, the conclusions from this paper supports the effectiveness of cluster formation policies (e.g. the creation of industrial parks), especially in a context in which the country is actively looking into the internationalization of its manufacturing industry. Second, if the country's export promotion policies are concerned not only on promoting the diversification of exporter's portfolio and supporting the entrance of new exporters, but also on making sure that the resulting new trade flows actually manage to survive for several years, then it might be worth to take extra care of firms located in areas with a lower concentration of exporters.

Finally, the results presented here can help to motivate future research on the role of spillovers on the export performance of firms. In particular, the recent availability of data that allows to identify the buyer behind each export transaction makes it a natural and interesting step to study whether these spillover effects are also be present when including the buyer dimension into the analysis.

References

Aeberhardt, R., Buono, I., and Fadinger, H. (2009). Learning, Incomplete Contracts and Export Dynamics: Theory and Evidence from French Firms. Vienna Economics Papers 1006, University of Vienna, Department of Economics.

Aitken, B., Hanson, G. H., and Harrison, A. E. (1997). Spillovers, Foreign Investment, and Export Behavior. *Journal of International Economics*, 43(1-2):103–132.

Albornoz, F., Fanelliz, S., and Hallak, J. C. (2013). Survival in export markets. Unpublished manuscript.

Albornoz-Crespo, F., Calvo Pardo, H. F., Corcos, G., and Ornelas, E. (2010). Sequential Exporting. CEPR Discussion Papers 8103, C.E.P.R. Discussion Papers.

Amador, J. and Opromolla, L. D. (2008). Product and Destination Mix in Export Markets. Working Papers w200817, Banco de Portugal, Economics and Research Department.

Araujo, L., Mion, G., and Ornelas, E. (2013). Institutions and export dynamics.

Bernard, A. B. and Jensen, J. B. (2004). Why some firms export. *The Review of Economics and Statistics*, 86(2):561–569.

Besedes, T. and Prusa, T. (2006). Ins, outs, and the duration of trade. *Canadian Journal of Economics*, 39(1):266–295.

Besedes, T. and Prusa, T. J. (2007). The role of extensive and intensive margins and export growth. NBER Working Papers 13628, National Bureau of Economic Research, Inc.

Buono, I., Fadinger, H., and Berger, S. (2008). The Micro Dynamic of Exporting-Evidence from French Firms. Vienna Economics Papers 0901, University of Vienna, Department of Economics.

Cadot, O., Iacovone, L., Pierola, D., and Rauch, F. (2011). Success and failure of african exporters. Policy Research Working Paper Series 5657, The World Bank.

Cameron, A. C. and Trivedi, P. K. (2010). *Microeconometrics Using Stata, Revised Edition*. Number musr in Stata Press books. StataCorp LP.

Cassey, A. and Schmeiser, K. (2012). The Agglomeration of Exporters by Destination. *The Annals of Regional Science*, pages 1–19.

Cebeci, T. (2012). A Concordance Among Harmonized System 1996, 2002 and 2007 Classifications. Technical report, World Bank.

Chaney, T. (2008). Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review*, 98(4):1707–21.

Cherkashin, I., Demidova, S., Kee, H. L., and Krishna, K. (2015). Firm heterogeneity and costly trade: A new estimation strategy and policy experiments. *Journal of International Economics*, 96(1):18–36.

Choquette, E. and Meinen, P. (2014). Export Spillovers: Opening the Black Box. *The World Economy*, pages n/a–n/a.

Clerides, S. K., Lach, S., and Tybout, J. R. (1998). Is Learning By Exporting Important? Micro-Dynamic Evidence From Colombia, Mexico, And Morocco. *The Quarterly Journal of Economics*, 113(3):903–947.

Crozet, M., Head, K., and Mayer, T. (2012). Quality Sorting and Trade: Firm-level Evidence for French Wine. *Review of Economic Studies*, 79(2):609–644.

Das, S., Roberts, M. J., and Tybout, J. R. (2007). Market entry costs, producer heterogeneity, and export dynamics. *Econometrica*, 75(3):837–873.

Dumont, M., Merlevede, B., Piette, C., and Rayp, G. (2010). The Productivity and Export Spillovers of the Internationalisation Behaviour of Belgian Firms. Working Paper Research 201, National Bank of Belgium.

Duranton, G. and Puga, D. (2004). Micro-Foundations of Urban Agglomeration Economies. In Henderson, J. V. and Thisse, J. F., editors, *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics*, chapter 48, pages 2063–2117. Elsevier.

Eaton, J., Eslava, M., Jinkins, D., Krizan, C. J., and Tybout, J. (2012). A Search and Learning Model of Export Dynamics. Technical report.

Eaton, J., Eslava, M., Kugler, M., and Tybout, J. (2007). Export Dynamics in Colombia: Firm-Level Evidence. NBER Working Papers 13531, National Bureau of Economic Research, Inc.

Eaton, J., Kortum, S. S., and Kramarz, F. (2009). An anatomy of international trade: Evidence from french firms. CEPR Discussion Papers 7111, C.E.P.R. Discussion Papers.

Esteve-Perez, S., Pallardo-Lopez, V., and Requena-Silvente, F. (2011). The duration of firm-destination export relationships: Evidence from spain, 1997-2006. Working Papers 1102, Department of Applied Economics II, Universidad de Valencia.

Fernandes, A. P. and Tang, H. (2014). Learning to Export from Neighbors. *Journal of International Economics*, 94(1):67–84.

Freund, C. and Pierola, M. D. (2010). Export Entrepreneurs : Evidence from Peru. Policy Research Working Paper Series 5407, The World Bank.

Fujita, M. and Thisse, J.-F. (2013). *Economics of Agglomeration. Cities, Industrial Location and Globalization*. Cambridge Books. Cambridge University Press, second edition edition.

Gorg, H., Kneller, R., and Murakozy, B. (2008). What makes a successful exporter? Discussion Papers 08/01, University of Nottingham, GEP.

Gormley, T. A. and Matsa, D. A. (2013). Common Errors: How to (and Not to) Control for Unobserved Heterogeneity. Unpublished manuscript.

Greenaway, D. and Kneller, R. (2008). Exporting, Productivity and Agglomeration. *European Economic Review*, 52(5):919–939.

Kang, Y. (2011). Is Agglomeration a Free Lunch for New Exporters? Evidence from Chile. Unpublish manuscrupt.

Koenig, P. (2009). Agglomeration and the Export Decisions of French Firms. *Journal of Urban Economics*, 66(3):186–195.

Koenig, P., Mayneris, F., and Poncet, S. (2010). Local Export Spillovers in France. *European Economic Review*, 54(4):622–641.

Lawless, M. and Whelan, K. (2008). Where Do Firms Export, How Much, and Why? Working Papers 200821, School Of Economics, University College Dublin.

Manova, K. and Zhang, Z. (2012). Export prices across firms and destinations. *The Quarterly Journal of Economics*, 127(1):379–436.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Melitz, M. J. and Ottaviano, G. I. P. (2008). Market Size, Trade, and Productivity. *Review of Economic Studies*, 75(1):295–316.

Molina, A. C. and Fugazza, M. (2009). The determinants of trade survival. IHEID Working Papers 05-2009, Economics Section, The Graduate Institute of International Studies.

Moral-Benito, E. (2013). Agglomeration Matters for Trade. 2013 Meeting Papers 85, Society for Economic Dynamics.

Morales, E., Sheu, G., and Zahler, A. (2015). Extended Gravity. Technical report, Princeton University.

Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit. *The Review of Economics and Statistics*, 72(2):334–38.

Paravisini, D., Rappoport, V., Schnabl, P., and Wolfenzon, D. (2011). Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. NBER Working Papers 16975, National Bureau of Economic Research, Inc.

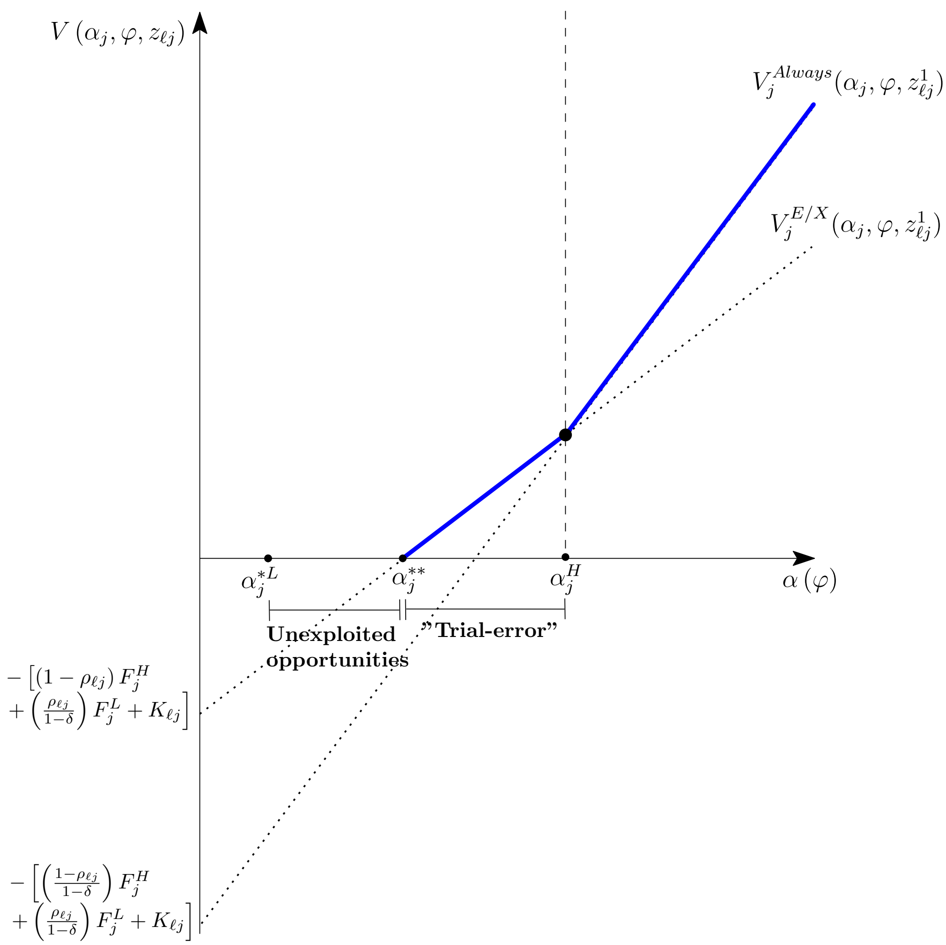
Rauch, J. E. and Watson, J. (2003). Starting Small in an Unfamiliar Environment. *International Journal of Industrial Organization*, 21(7):1021–1042.

Roberts, M. J. and Tybout, J. R. (1995). An Empirical Model of Sunk Costs and the Decision to Export. Policy Research Working Paper Series 1436, The World Bank.

Segura-Cayuela, R. and Vilarrubia, J. M. (2008). Uncertainty and Entry into Export Markets. Banco de Espaa Working Papers 0811, Banco de Espaa.

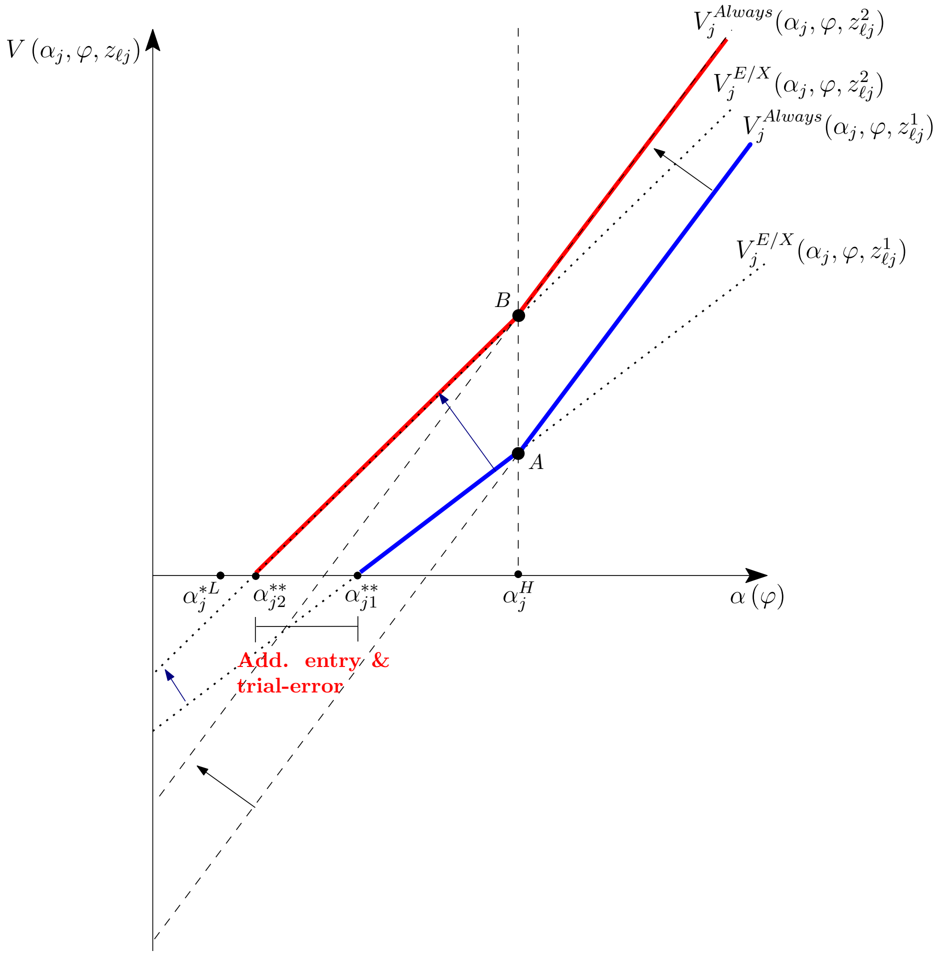
Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, 68(1):115–132.

Figure 1
Ex-ante and ex-post cutoffs



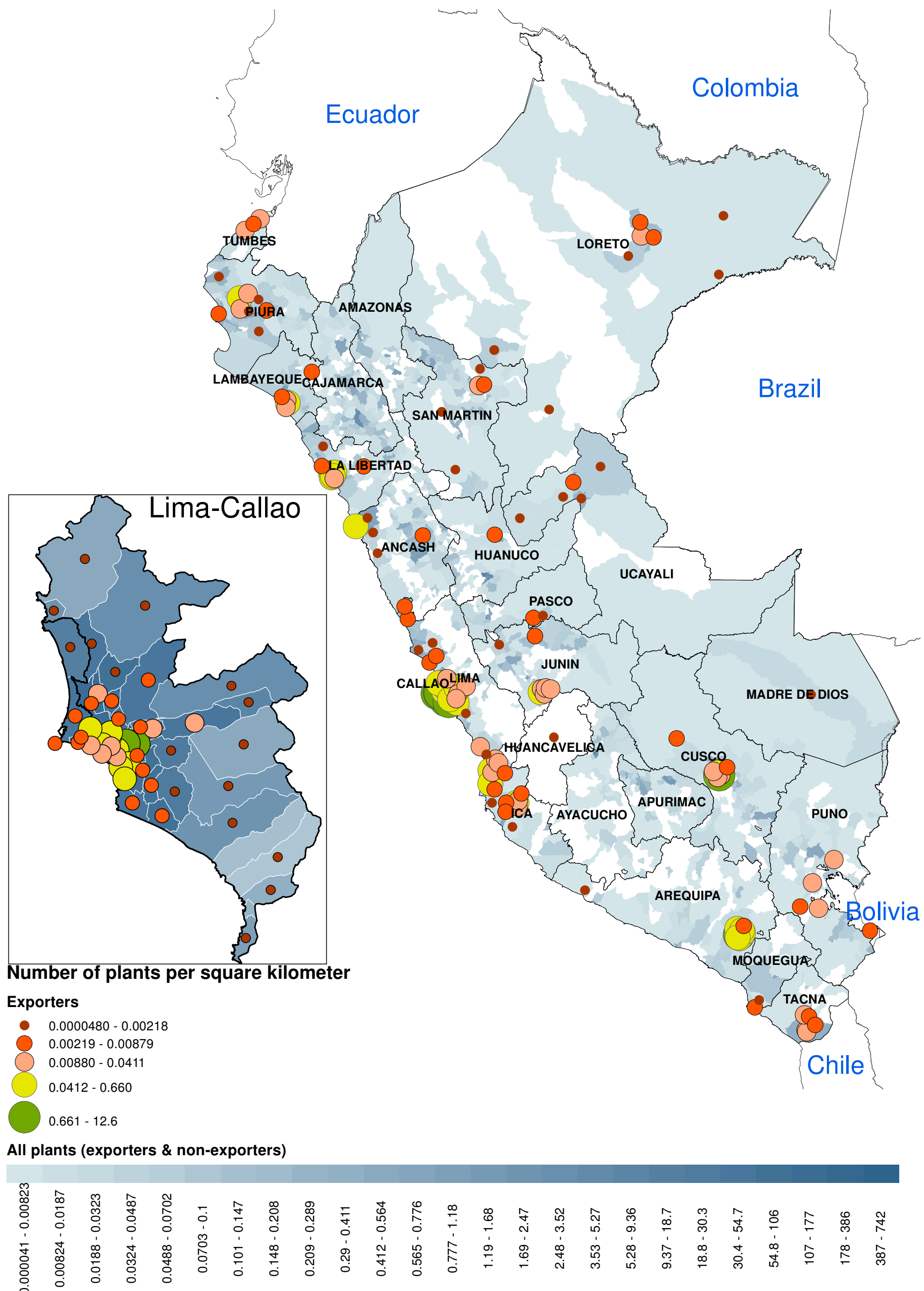
Notes: This figure illustrates the value of exporting to country j for a firm with productivity $\varphi > \varphi*$. V_j^{Always} is the value function for firms that always enter to j. $V_j^{E/X}$ is the value function of firms that follow the entry-exit strategy: the will continue to export to j only if they draw a low fixed cost realization. The ex-ante cutoff α_j^{**} is defined in (6), while the ex-post cutoff is defined in (3). The cutoff α_j^{*L} is the relevant cutoff defined in (2)when there is no uncertainty, and when $K_j = K(z_{\ell j})$ and $F_j = F_j^L$.

Figure 2
Effects of an increase in $z_{\ell j,t-1}$



Notes: This figure illustrates the effects on an increase in $z_{\ell j}$. The distance between α_{j1}^{**} and α_{j2}^{**} represents the additional measure of firms that now are able to exploit the otherwise unexploited opportunities. These firms engage in the entry-exit strategy and so they increase the rate of entry.

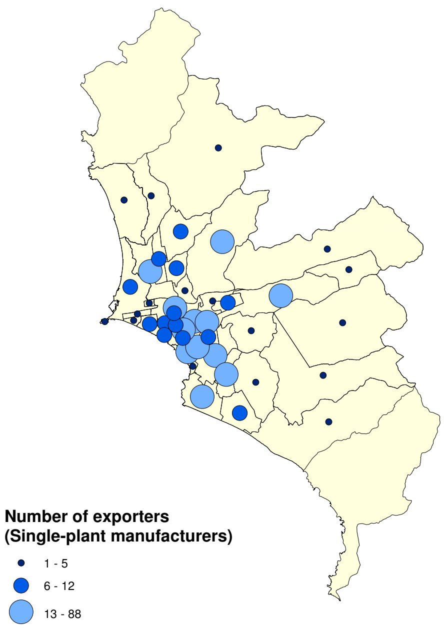
Figure 3
Location of single-plant exporters of manufactured products, Peru 2006



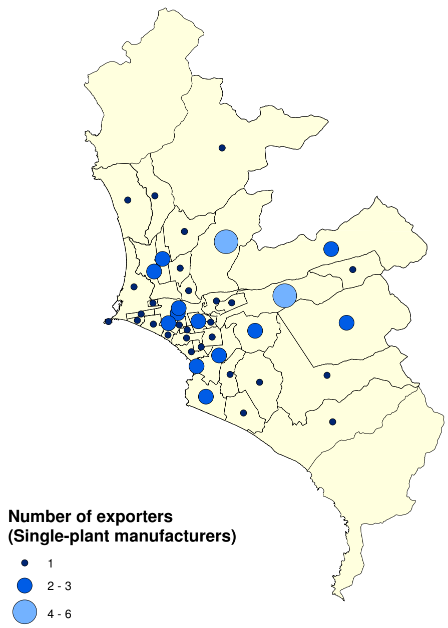
Notes: The color-ramp is used to denote the agglomeration of plants at the district level; darker shades imply higher agglomeration. The source of this data is the National Registry of Manufactured firms constructed by the Ministry of Production, based on the registry for the National Economic Census conducted in 2007. The colored circles denote bins for the number of single-plant exporting manufacturing firms in 2006. Areas delimited by black lines correspond to *regiones*.

Figure 4
Location of single-plant exporters of manufactured products, selected industries, Province of Lima 2006

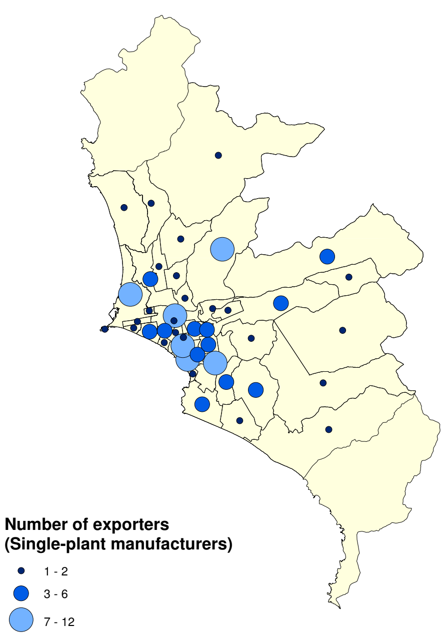
(a) Textiles



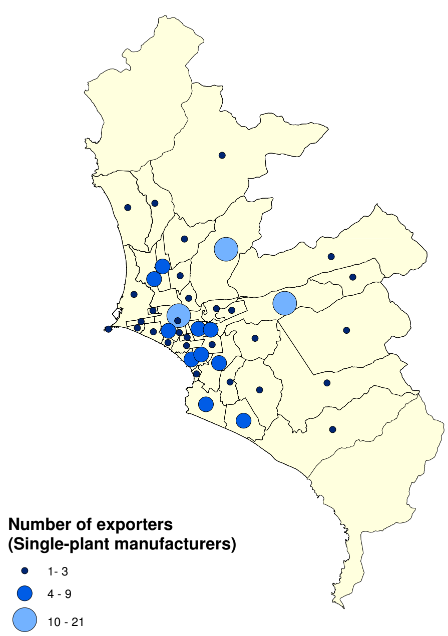
(b) Ceramic products



(c) Furniture



(d) Plastic & rubber products



Notes: Bins denote terciles of the number of firms in each industry.

Table 1
Summary statistics, selected years

	2002	2003	2004	2005	2006	2007
Aggregate export sales (USD mill.)	1,260	1,394	1,774	2,106	2,486	2,873
# of firms	1,260	1,377	1,523	1,695	1,828	1,914
# of products (HS6)	470	467	483	502	497	497
# of products (HS2)	73	74	72	72	73	72
# of destinations	110	127	141	145	141	142

Notes: All product codes are consistent across years using the concordance tables provided by Cebeci 2012.

Table 2
Distribution of firms in sample according to their number of plants

Number of plants	2002	2003	2004	2005	2006	2007
1	89.3	89.9	90.9	91.4	91.4	91.2
2	5.1	5.0	4.7	4.6	4.4	4.6
3	1.5	1.3	1.3	1.1	1.4	1.3
4	1.4	1.2	1.1	1.1	1.0	1.0
5-10	1.9	1.8	1.4	1.2	1.2	1.4
11+	0.8	0.8	0.8	0.6	0.5	0.5
Total number of exporters	1,260	1,377	1,523	1,695	1,828	1,913
Single plant exporters	1,098	1,122	1,235	1,382	1,548	1,669
in Lima	838	899	966	1,095	1,205	1,305
in other regions	260	223	269	287	343	364

Notes: The sample consists of manufacturing firms that registered exports between 2002 and 2007. Exporters are classified as single-plant or multi-plant firms according to the number of establishments declared as "plants" before the local tax authority.

Table 3
Local export spillovers on entry

	(1)	(2)	(3)	(4)	(5)	(6)
# of manufacturers; same prod. & cty	0.629***	0.579***	0.420***			0.399***
	(0.093)	(0.083)	(0.037)			(0.033)
# of manufacturers; same prod.				0.007***		
				(0.001)		
# of manufacturers; same country					0.017**	
					(0.007)	
# of manufacturers; same prod to other cty						0.008
						(0.006)
# of manufacturers; other prod to same cty						0.000
						(0.001)
log(1+dest. country's product specific imports)		0.034***	0.022***	0.028***	0.021***	0.019***
		(0.008)	(0.005)	(0.005)	(0.004)	(0.003)
log(distance)		-0.140***	-0.087***	-0.110***	-0.083***	-0.076***
		(0.033)	(0.022)	(0.022)	(0.018)	(0.018)
Common border		0.101***	0.038	0.062**	-0.068*	-0.021
		(0.037)	(0.023)	(0.025)	(0.039)	(0.038)
Common language		-0.009	-0.028	-0.023	-0.058***	-0.044***
		(0.024)	(0.019)	(0.019)	(0.017)	(0.015)
# of HS2 codes exported to same country			0.573***	0.590***	0.566***	0.563***
			(0.077)	(0.076)	(0.072)	(0.073)
# of other countries to which prod. is exported			0.036***	0.035***	0.036***	0.036***
			(0.004)	(0.004)	(0.004)	(0.004)
Firm-product extended gravity index			3.891***	3.942***	3.938***	3.893***
			(1.103)	(1.110)	(1.108)	(1.104)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes
Num. obs	11,142,803	11,142,803	11,142,803	11,142,803	11,142,803	11,142,803
Adjusted R2	0.020	0.021	0.057	0.056	0.056	0.057
Num. of clusters	166	166	166	166	166	166

Notes: The dependent variable is a binary indicator taking the value of 1 if a manufacturer started to export product p to country d at time t as in equation (8). The sample consists of all new export relationships initiated between 2002 and 2007 by single-plant manufacturers. Spillovers at the firm-product definition are calculated as the number of firms exporting the same 4-digit HS product code unless otherwise stated. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district level. All count regressors are lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4

Heterogeneity, spatial decay and time decay on entry

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
# of manufacturers; same prod. & cty	0.420***	0.803***	0.309***	0.496***		0.178***	
	(0.037)	(0.222)	(0.032)	(0.056)		(0.038)	
log(firm's total exports) × # of manufacturers; same prod. & cty		-0.035*					
		(0.019)					
(young exporter=1) × # of manufacturers; same prod. & cty			0.484***				
			(0.166)				
# of manufacturers; same prod. & cty squared				-0.007***			
				(0.003)			
# of manufacturers in province; same prod. & country					0.049***		
					(0.005)		
# of manufacturers in other districts in prov; same prod. & cty						0.041***	
						(0.005)	
# of new manufacturers; same prod. & country							0.627***
							(0.143)
# of old manufacturers; same prod. & country							0.307***
							(0.048)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	yes	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes	yes
Num. obs	11,142,803	11,142,803	11,142,803	11,142,803	11,142,803	11,142,803	11,142,803
Adjusted R2	0.057	0.057	0.057	0.057	0.058	0.058	0.057
Num. of clusters	166	166	166	166	166	166	166

Notes: The dependent variable is a binary indicator taking the value of 1 if a manufacturer started to export product p to country d at time t as in equation (8). The sample consists of all new export relationships initiated between 2002 and 2007 by single-plant manufacturers. Spillovers at the firm-product definition are calculated as the number of firms exporting the same 4-digit HS product code unless otherwise stated. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district level. All count regressors are lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5
Local export spillovers on survival

	(1)	(2)	(3)	(4)	(5)	(6)
# of manufacturers; same prod. & cty	1.609***	1.391***	1.171***			1.143***
	(0.467)	(0.435)	(0.408)			(0.424)
# of manufacturers; same prod.				0.104		
				(0.143)		
# of manufacturers; same country					0.121***	
					(0.034)	
# of manufacturers; same prod to other cty						0.037
						(0.044)
# of manufacturers; other prod to same cty						-0.036
						(0.102)
log(1+dest. country's product specific imports)		0.713***	0.572***	0.687***	0.579***	0.541***
		(0.193)	(0.163)	(0.159)	(0.163)	(0.164)
log(distance)		-4.680***	-3.753**	-4.628***	-3.542**	-3.440**
		(1.701)	(1.562)	(1.737)	(1.786)	(1.729)
Common border		0.087	1.575	0.823	0.902	1.583
		(2.492)	(2.258)	(2.434)	(2.547)	(2.290)
Common language		0.268	-1.405	-2.049	-1.215	-1.164
		(1.887)	(1.817)	(1.644)	(1.681)	(1.728)
# of HS2 codes exported to same country			1.066***	1.087***	1.035***	1.050***
			(0.135)	(0.141)	(0.142)	(0.144)
# of other countries to which prod. is exported			0.185	0.167	0.184	0.190
			(0.117)	(0.117)	(0.115)	(0.116)
Firm-product extended gravity index			1.782***	1.741***	1.788***	1.796***
			(0.269)	(0.272)	(0.269)	(0.270)
logX			4.460***	4.496***	4.492***	4.460***
			(0.159)	(0.157)	(0.156)	(0.160)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes
Num. obs	24,579	24,579	24,579	24,579	24,579	24,579
Adjusted R2	0.364	0.366	0.402	0.401	0.401	0.402

Notes: The dependent variable is a binary indicator taking the value of 1 if the export relationship initiated by a manufacturer started to export product p to country d at time t survived beyond its k year(s) (i.e. lasted at least $k+1$ years) as defined in (9). The sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district level. All regressors and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6
Heterogeneity, spatial decay and time decay on survival

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
# of manufacturers; same prod. & cty	1.171***	0.241	0.940*	1.759***		0.681	
	(0.408)	(0.649)	(0.505)	(0.486)		(0.433)	
log(firm's total exports) × # of manufacturers; same prod. & cty		0.082*					
		(0.049)					
(young exporter=1) × # of manufacturers; same prod. & cty			0.788				
			(0.483)				
# of manufacturers; same prod. & cty squared				-0.037***			
				(0.013)			
# of manufacturers in province; same prod. & country					0.132***		
					(0.032)		
# of manufacturers in other districts in prov; same prod. & cty						0.098***	
						(0.030)	
# of new manufacturers; same prod. & country							0.381
							(0.505)
# of old manufacturers; same prod. & country							1.690***
							(0.311)
(1 if manufacturer is new)=1 × # of manufacturers; same prod. & cty							
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	yes	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes	yes
Num. obs	24,579	24,579	24,579	24,579	24,579	24,579	24,579
Adjusted R2	0.402	0.402	0.402	0.402	0.403	0.403	0.402
Num. of clusters	162	162	162	162	162	162	162

Notes: The dependent variable is a binary indicator taking the value of 1 if the export relationship initiated by a manufacturer started to export product p to country d at time t survived beyond its k year(s) (i.e. lasted at least $k + 1$ years) as defined in (9). The sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district level. All regressors and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix

A Proofs of Propositions

Proof of Proposition 1. To show that the ex-ante cutoff $\alpha^{**}_{\ell j,t}$ is decreasing in $z_{\ell j}$ denote by A and B the numerator and denominator in (6). We need to show that the sign of the derivative $\partial\alpha^{**}_j(\varphi, z_{\ell j,t-1})/\partial z_{\ell j,t-1} = (A'B - B'A)/B^2$ is negative. Applying the derivatives and after some rearrangements we get:

$$\frac{\partial\alpha^{**}_j(\varphi, z_{\ell j})}{\partial z_{\ell j}} = \frac{\varphi^{\sigma-1}\Lambda_j\left[(1-\delta)(1-\delta+\delta\rho)K'_{\ell j} - \rho'\delta(1-\delta)F^H_j - \rho'(1-\delta)\left(F^H_j - F^L_j\right)\right]}{\left[(1-\delta+\delta\rho)\varphi^{\sigma-1}\Lambda_j\right]^2}$$

which is strictly negative because the term in brackets in the numerator is unambiguously negative given the assumption that $K'_{\ell j,t} \leq 0$.

Proof of Proposition 2. The proportion of firms that drew a demand shock between $\alpha^*_{\ell j}$ and α^H_j is given by

$$\xi^H_{\ell j} \equiv \frac{\int_{\varphi^*_j}^{+\infty}\int_{\alpha^{**}_j(\varphi,z_{\ell j})}^{\alpha^H_j(\varphi)} dH_j(\alpha)\,dG(\varphi)}{\int_{\varphi^*_j}^{+\infty}\int_{\alpha^{**}_j(\varphi,z_{\ell j})}^{+\infty} dH_j(\alpha)\,dG(\varphi)}$$

where $\bar{\varphi}$ and $\bar{\alpha}$ are the upper-bounds of support of the cumulative distributions $G(\cdot)$ and $H(\cdot)$ respectively.

To show that this proportion is increasing in $z_{\ell j}$, denote by A and B the numerator and the denominator of $\xi^H_{\ell j}$ and write:

$$A\left(\alpha^{**}_j\right) = \int_{\varphi^*_j}^{+\infty}\left[H\left(\alpha^H_j(\varphi)\right) - H\left(\alpha^{**}_j(\varphi, z_{\ell j})\right)\right]dG(\varphi) \tag{11}$$

$$B\left(\alpha^{**}_j\right) = \int_{\varphi^*_j}^{+\infty}\left[1 - H\left(\alpha^{**}_j(\varphi, z_{\ell j})\right)\right]dG(\varphi) \tag{12}$$

Then, applying Leibniz's rule for differention under the integral sign it can be shown that:

$$\frac{\partial A\left(\alpha^{**}_j\right)}{\partial z_{\ell j}} = -\int_{\varphi^*_j}^{+\infty}\left[\frac{\partial H\left(\alpha^{**}_j\right)}{\partial\alpha^{**}_j}\frac{\partial\alpha^{**}_j}{\partial z_{\ell j}}\right]dG(\varphi) = \frac{\partial B\left(\alpha^{**}_j\right)}{\partial z_{\ell j}}$$

Finally, the derivative of $\xi^H_{\ell j}$ with respect to $z_{\ell j}$ is:

$$
\begin{aligned}
\frac{\partial \xi^H_{\ell j}}{\partial z_{\ell j}} &= \frac{1}{B\left(\alpha^{**}_j\right)^2}\left[\frac{\partial A\left(\alpha^{**}_j\right)}{\partial z_{\ell j}} B\left(\alpha^{**}_j\right) - A\left(\alpha^{**}_j\right)\frac{\partial B\left(\alpha^{**}_j\right)}{\partial z_{\ell j}}\right] \\
&= \frac{1}{B\left(\alpha^{**}_j\right)^2}\left[B\left(\alpha^{**}_j\right) - A\left(\alpha^{**}_j\right)\right]\left[\frac{\partial A\left(\alpha^{**}_j\right)}{\partial z_{\ell j}}\right]
\end{aligned}
$$

which will be positive because $B\left(\alpha^{**}_j\right) > A\left(\alpha^{**}_j\right)$ as inferred from (11) and (12).

Proof of Proposition 3. The overall survival rate among entrants will be a weighted sum of the survival rates of the two types of firms entering in a given period:

$$P\left(Survival_{\ell j}\left(\varphi, \alpha, z_{\ell j}\right) = 1 \,\middle|\, \varphi \geq \varphi^*_j\right) = \left(1 - \rho_{\ell j}\right)\cdot\left(1 - \xi^H_{\ell j}\right) + \rho_{\ell j}\cdot 1 \tag{13}$$

To see how this survival rate changes when $z_{\ell j}$ increases we need to analyse the sign of the derivative:

$$\frac{\partial P\left(Survival_{\ell j}\left(\varphi, \alpha, z_{\ell j}\right) = 1 \,\middle|\, \varphi \geq \varphi^*_j\right)}{\partial z_{\ell j}} = \frac{\partial\left(1 - \rho_{\ell j}\right)\cdot\left(1 - \xi^H_{\ell j}\right)}{\partial z_{\ell j}} + \frac{\partial \rho_{\ell j}}{\partial z_{\ell j}} \tag{14}$$

By assumption, the last derivative $\rho' \equiv \partial\rho/\partial z_{\ell j}$ is positive. We can then show that the sign of the first term will be unambiguously negative

$$
\begin{aligned}
\frac{\partial\left(1 - \rho_{\ell j}\right)\cdot\left(1 - \xi^H_{\ell j}\right)}{\partial z_{\ell j}} &= -\rho^{'}_{\ell j} - \xi^{'H}_{\ell j,t} + \rho_{\ell j}\xi^{'H}_{\ell j} + \rho^{'}_{\ell j}\xi^H_{\ell j} \\
&= -\left[\rho'_{\ell j}\left(1 - \xi^H_{\ell j}\right) + \xi^{'H}_{\ell j}\left(1 - \rho_{\ell j}\right)\right]
\end{aligned}
$$

because all terms in brackets are positive. Then it is evident that the sign of the whole derivative (14) will depend upon which derivative dominates in the right-hand side. Given the assumptions made on the function $\rho_{\ell j}$, the derivative $\partial\rho_{\ell j}/\partial z_{\ell j}$ will be positive and large when $z_{\ell j}$ is relatively low and so in that case the overall survival rate of entrants in region ℓ will be higher. Otherwise, if $z_{\ell j}$ raises while being too high, then the derivative $\partial\rho_{\ell j}/\partial z_{\ell j}$ will be low and hence survival will be end up being lower.

B Data Appendix

B.1 Data cleaning process

Because the aim of this paper was the study of trade relationships defined as the firm-product-country combination , several challenges that are endemic to the work with customs data had to be taken into account. In customs data the entity doing the international trade operation is not necessarily a for-profit business in the usual way we understand it. Besides the usual presence of transactions carried by entities such as public institutions, airlines, ship lines, embassies, and individuals sending taxable goods, there are other transactions carried by regular firms that should not be taken into account when studying export flows.

First, there were firms selling products that were not part of their main business activity and that could be identified as occasional exports for specific purposes. For example, there are a number of shipments sent by mining companies, larger beer brewers and large food companies of apparel products. While these shipments could in principle be the result of an aggressive horizontal diversification process, they are most likely the result of occasional exports of gifts oriented at promoting firms in international fairs or events of the like. The same inconsistency was found with leather products, and coffee products. Second, there are a number of also occasional transactions of capital goods by firms from several industries. These shipments are due to the selling abroad of movable fixed assets as part of the liquidation process of the domestic operation of the firm.

The first problem can be ameliorated by contrasting the firm's main activity line as stated by their self-declared industrial codes (ISIC rev. 3) and the correlated ISIC codes on the products being sold. There were a few major discrepancies which were dropped from the analysis. The second problem is solved by focusing on exports of consumption goods and intermediate goods only. While there are some Peruvian manufacturers that do export items such as machinery, they represent only a small share of export transactions.

Another problem in using customs data are the changes in the product classification. The Peruvian data uses the NANDINA product classification common to all members of the Andean Community (Bolivia, Colombia, Ecuador, Peru and Venezuela), whose 6 first digits correspond to the Harmonized System. Besides major revisions to the HS classification system, the most recent of which occurred in 1992, 1996, 2002, 2007 and recently in 2012 there are a number of on-the-go revisions that can occur in any year and that include both merging and splitting of product codes. To deal with these changes in classification, I used the concordance developed by Cebeci (2012) for the HS classification system.

C Additional figures and tables

Table A1
Distribution of dependent variable for entry according to the number of neighbors, year 2006

Number of neighbors exporting...	Entry=0 Triads	Entry=0 Firms	Entry=1 Triads	Entry=1 Firms
...same product, same country				
0	2,396,482	1,459	3,805	941
1 - 5	126,081	1,367	1,955	571
6 - 10	4,620	606	290	169
11 - 20	1,338	345	116	71
21 - 50	527	183	105	45
...same product, all countries				
0	61,953	836	1,033	445
1 - 5	1,653,172	1,423	2,687	770
6 - 10	375,595	932	955	372
11 - 20	254,514	697	771	313
21 - 50	149,675	439	609	205
50+	34,139	144	216	85
...all products, same country				
0	1,101,688	1,441	709	216
1 - 5	1,008,345	1,429	2,118	505
6 - 10	177,961	1,179	911	264
11 - 20	125,868	1,059	1,014	278
21 - 50	85,571	802	864	275
50+	29,615	495	655	207

Notes: This table shows the distribution of the dependent variable for entry defined in (7) according to the number of neighbors for manufacturing firms that registered exports in 2006. A triad corresponds to a firm-product-country combination.

Table A2

Distribution of dependent variable for survival according to the number of neighbors, year 2006

Number of neighbors exporting...	Survival=0 Triads	Survival=0 Firms	Survival=1 Triads	Survival=1 Firms
...same product, same country				
0	8,211	2,036	2,537	533
1 - 5	4,981	1,198	9,505	1,706
6 - 10	604	292	1,006	444
11 - 20	259	147	472	234
21 - 50	126	61	340	132
...same product, all countries				
0	2,419	1,002	757	215
1 - 5	6,236	1,577	6,519	1,377
6 - 10	2,259	772	2,399	690
11 - 20	1,831	661	2,059	596
21 - 50	1,207	378	1,725	464
50+	229	98	401	132
...all products, same country				
0	1,279	364	437	102
1 - 5	5,108	978	4,818	848
6 - 10	1,956	509	2,009	434
11 - 20	2,061	525	2,227	521
21 - 50	2,330	608	2,611	544
50+	1,447	427	1,758	411

Notes: This table shows the distribution of the dependent variable for survival defined in (9) according to the number of neighbors for manufacturing firms that registered exports in 2006. A triad corresponds to a firm-product-country combination.

Table A3

Robustness checks on entry: Spillover measures adjusted for area

	(1)	(2)	(3)	(4)	(5)	(6)
# of manufacturers; same prod. & cty	7.568***	7.152***	6.310***			5.429***
	(0.638)	(0.581)	(0.638)			(0.671)
# of manufacturers; same prod.				0.044***		
				(0.009)		
# of manufacturers; same country					0.883***	
					(0.075)	
# of manufacturers; same prod to other cty						0.550***
						(0.107)
# of manufacturers; other prod to same cty						-0.040***
						(0.013)
log(1+dest. country's product specific imports)		0.041***	0.031***	0.037***	0.022***	0.023***
		(0.008)	(0.005)	(0.006)	(0.005)	(0.005)
log(distance)		-0.173***	-0.127***	-0.155***	-0.093***	-0.095***
		(0.035)	(0.027)	(0.031)	(0.026)	(0.025)
Common border		0.124**	0.068**	0.066*	-0.129**	-0.054
		(0.048)	(0.031)	(0.035)	(0.050)	(0.038)
Common language		0.002	-0.015	-0.015	-0.058**	-0.041*
		(0.028)	(0.022)	(0.028)	(0.024)	(0.024)
# of HS2 codes exported to same country			0.566***	0.583***	0.529***	0.536***
			(0.075)	(0.076)	(0.074)	(0.075)
# of other countries to which prod. is exported			0.034***	0.033***	0.034***	0.034***
			(0.004)	(0.004)	(0.004)	(0.004)
Firm-product extended gravity index			3.894***	3.938***	3.936***	3.902***
			(1.094)	(1.108)	(1.105)	(1.095)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes
Num. obs	13,786,228	13,786,228	13,786,228	13,786,228	13,786,228	13,786,228
Adjusted R2	0.029	0.030	0.056	0.053	0.055	0.057

Notes: The dependent variable is a binary indicator taking the value of 1 if a manufacturer started to export product p to country d at time t as in equation (8). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 3 are included but not reported for space considerations. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district-market level. All count regressors are in logarithms and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4

Robustness checks on entry: volume of exports as spillover measure

	(1)	(2)	(3)	(4)	(5)	(6)
log(manufs' exports from district-same prod. & country)	0.203***	0.185***	0.127***			0.118***
	(0.031)	(0.029)	(0.022)			(0.020)
log(manufacturers' exports from district-same prod.)				-0.006***		
				(0.002)		
log(manufacturers' exports from district-same country)					0.028***	
					(0.006)	
log(exports from district; other prod-same country)						-0.012***
						(0.002)
log(exports from district; same prod-other countries)						0.022***
						(0.004)
log(1+dest. country's product specific imports)		0.036***	0.029***	0.037***	0.021***	0.017***
		(0.008)	(0.006)	(0.006)	(0.004)	(0.004)
log(distance)		-0.167***	-0.130***	-0.155***	-0.124***	-0.107***
		(0.035)	(0.028)	(0.031)	(0.031)	(0.028)
Common border		0.031	0.004	0.066*	-0.028	-0.066
		(0.046)	(0.037)	(0.035)	(0.052)	(0.043)
Common language		-0.037	-0.041	-0.015	-0.119***	-0.120***
		(0.024)	(0.026)	(0.028)	(0.028)	(0.026)
# of HS2 codes exported to same country			0.562***	0.583***	0.558***	0.543***
			(0.075)	(0.076)	(0.071)	(0.071)
# of other countries to which prod. is exported			0.034***	0.035***	0.034***	0.038***
			(0.004)	(0.004)	(0.004)	(0.004)
Firm-product extended gravity index			3.825***	3.938***	3.929***	3.829***
			(1.083)	(1.108)	(1.103)	(1.084)
Firm-year FE	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes
N	13,786,228	13,786,228	13,786,228	13,786,228	13,786,228	13,786,228

Notes: The dependent variable is a binary indicator taking the value of 1 if a manufacturer started to export product p to country d at time t as in equation (8). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 3 are included but not reported for space considerations. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district-market level. All count regressors are in logarithms and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A5

Robustness checks on survival: Spillover measures adjusted for area

	(1)	(2)	(3)	(4)	(5)	(6)
# of manufacturers; same prod. & cty	9.234**	7.319**	5.891*			5.772*
	(4.015)	(3.513)	(3.227)			(3.173)
# of manufacturers; same prod.				0.027		
				(1.240)		
# of manufacturers; same country					1.437*	
					(0.823)	
# of manufacturers; same prod to other cty						0.709
						(0.899)
# of manufacturers; other prod to same cty						-0.515
						(1.069)
Firm-year FE	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes
Num. obs	24,579	24,579	24,579	24,579	24,579	24,579
Adjusted R2	0.362	0.365	0.401	0.401	0.401	0.401

Notes: The dependent variable is a binary indicator taking the value of 1 if the export relationship initiated by a manufacturer started to export product p to country d at time t survived beyond its k year(s) (i.e. lasted at least $k+1$ years) as defined in (9). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 5 are included but not reported for space considerations. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district level. All count regressors are in logarithms and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A6

Robustness checks on survival: volume of exports as spillover measure

	(1)	(2)	(3)	(4)	(5)	(6)
log(manufs' exports from district; same prod. & cty)	0.822***	0.695***	0.527***			0.404***
	(0.116)	(0.117)	(0.117)			(0.129)
log(manufacturers' exports from district; same prod.)				0.147		
				(0.090)		
log(manufacturers' exports from district; same cty)					0.617***	
					(0.147)	
log(manufs' exports from district; other prod to same cty)						0.123
						(0.096)
log(manufs' exports from district; same prod to other ctys)						0.542***
						(0.151)
Firm-year FE	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes
Num. obs	24,579	24,579	24,579	24,579	24,579	24,579
Adjusted R2	0.365	0.366	0.402	0.401	0.403	0.403

Notes: The dependent variable is a binary indicator taking the value of 1 if the export relationship initiated by a manufacturer started to export product p to country d at time t survived beyond its k year(s) (i.e. lasted at least $k+1$ years) as defined in (9). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 5 are included but not reported for space considerations. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district level. All count regressors are in logarithms and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A7

Robustness checks on entry: Different sub-samples of exporters

	(1) $nw < 10$	(2) $nw \geq 10$	(3) $nw < 50$	(4) $nw \geq 50$	(5) $3 \leq nw \leq 250$	(6) Lima	(7) Provinces	(8) Whole
# of manufacturers; same prod. & cty	0.744***	0.687***	0.672***	0.724***	0.598***	0.718***	1.034	0.718***
	(0.103)	(0.109)	(0.097)	(0.120)	(0.062)	(0.104)	(1.143)	(0.107)
prov=1 × # of manufacturers; same prod. & cty								0.094
								(1.652)
log(1+dest. country's product specific imports)	0.030***	0.024***	0.022***	0.028***	0.019***	0.025***	0.103***	0.026***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.004)	(0.005)	(0.031)	(0.005)
log(distance)	-0.143***	-0.096***	-0.114***	-0.102***	-0.092***	-0.108***	-0.217	-0.108***
	(0.029)	(0.027)	(0.027)	(0.028)	(0.022)	(0.026)	(0.271)	(0.026)
Common border	-0.082	0.086**	0.050	0.032	0.120***	0.014	0.553	0.033
	(0.053)	(0.035)	(0.040)	(0.048)	(0.025)	(0.022)	(0.537)	(0.029)
Common language	-0.030	-0.016	-0.007	-0.037	-0.014	-0.023	-0.034	-0.023
	(0.042)	(0.020)	(0.026)	(0.028)	(0.019)	(0.022)	(0.299)	(0.023)
# of HS2 codes exported to same country	0.597***	0.524***	0.650***	0.495***	0.525***	0.524***	0.821***	0.555***
	(0.101)	(0.089)	(0.088)	(0.094)	(0.102)	(0.075)	(0.300)	(0.075)
# of other countries to which prod. is exported	0.034**	0.037***	0.058***	0.041***	0.033***	0.039***	0.007	0.034***
	(0.016)	(0.004)	(0.018)	(0.004)	(0.007)	(0.004)	(0.030)	(0.004)
Firm-product extended gravity index	18.291***	3.479***	14.045***	2.917***	5.016***	4.118***	2.538***	3.852***
	(3.501)	(0.971)	(2.297)	(0.825)	(1.496)	(1.410)	(0.829)	(1.092)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	yes	yes	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes	yes	yes
Num. obs	3,460,206	10,326,022	6,470,514	7,315,714	7,721,263	13,500,014	286,214	13,786,228
Adjusted R2	0.087	0.058	0.090	0.056	0.063	0.045	0.139	0.058

Notes: The dependent variable is a binary indicator taking the value of 1 if a manufacturer started to export product p to country d at time t as in equation (8). Sample consists of all new export relationships initiated between 2002 and 2007 by single-plant manufacturers. Spillovers at the firm-product definition are calculated as the number of firms exporting the same 4-digit HS product code unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 3 are included but not reported for space considerations. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district level. All count regressors are lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A8

Robustness checks on survival: Different subsamples of exporters

	(1) $nw < 10$	(2) $nw \geq 10$	(3) $nw < 50$	(4) $nw \geq 50$	(5) $3 \leq nw \leq 250$	(6) Lima	(7) Provinces	(8) Whole
# of manufacturers; same prod. & cty	1.289***	1.123**	1.149***	1.181**	1.182**	1.112**	4.371**	1.151***
	(0.257)	(0.510)	(0.272)	(0.595)	(0.503)	(0.420)	(1.698)	(0.409)
(prov=1) × # of manufacturers; same prod. & cty								3.672
								(2.806)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	yes	yes	yes	yes	yes	yes	yes	yes
Prod. FE	yes	yes	yes	yes	yes	yes	yes	yes
District. FE	yes	yes	yes	yes	yes	yes	yes	yes
Num. obs	5,629	18,950	10,427	14,152	13,241	21,390	3,189	24,579
Adjusted R2	0.429	0.393	0.438	0.372	0.418	0.408	0.367	0.402

Notes: The dependent variable is a binary indicator taking the value of 1 if the export relationship initiated by a manufacturer started to export product p to country d at time t survived beyond its k year(s) (i.e. lasted at least $k+1$ years) as defined in (9). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 5 are included but not reported for space considerations. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district-market level. All count regressors are in logarithms and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A9
Robustness checks on survival: Stricter survival definitions

	(1) $Surv_t^{k=2}$	(2) $Surv_t^{k=3}$
# of manufacturers; same prod. & cty	0.851	0.682**
	(0.559)	(0.292)
Spill. def.	# of firms	# of firms
Firm-year FE	yes	yes
Prod. FE	yes	yes
District. FE	yes	yes
Num. obs	24,579	24,579
Adjusted R2	0.398	0.391

Notes: The dependent variable is a binary indicator taking the value of 1 if the export relationship initiated by a manufacturer started to export product p to country d at time t survived beyond its k year(s) (i.e. lasted at least $k+1$ years) as defined in (9). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 5 are included but not reported for space considerations. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district-market level. All count regressors are in logarithms and lagged one period. Significance levels: $^{***}\ p < 0.01$, $^{**}\ p < 0.05$, $^{*}\ p < 0.1$

Table A10

Robustness checks on entry: Alternative fixed effects specifications

	(1) Entry	(2) Entry	(3) Entry	(4) Entry	(5) Entry
# of manufacturers; same prod. & cty	0.420***	0.434***	0.293***	0.450***	0.420***
	(0.037)	(0.039)	(0.039)	(0.043)	(0.037)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	Yes	Yes	Yes	No	No
Firm FE	No	No	No	Yes	Yes
Year FE	No	No	No	No	No
Product FE	Yes	No	Yes	No	Yes
District FE	Yes	No	No	No	No
District-year FE	No	No	No	No	Yes
HS2-district FE	No	Yes	No	No	No
Country-district FE	No	No	Yes	No	No
Product-district-year FE	No	No	No	Yes	No
Num. obs	11,142,803	11,142,803	11,142,803	11,142,803	11,142,803
Adjusted R2	0.057	0.061	0.063	0.172	0.054

Notes: The dependent variable is a binary indicator taking the value of 1 if a manufacturer started to export product p to country d at time t as in equation (8). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 3 are included but not reported for space considerations. Specifications that do not include firm-year fixed effects include the firm's lagged total export sales as a proxy for total factor productivity to control for firm-level heterogeneity. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district-market level. All count regressors are in logarithms and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A11

Robustness checks on survival: Alternative fixed effects specifications

	(1) $Surv_t^{k=1}$	(2) $Surv_t^{k=1}$	(3) $Surv_t^{k=1}$	(4) $Surv_t^{k=1}$	(5) $Surv_t^{k=1}$
# of manufacturers; same prod. & cty	1.171***	1.384***	0.967*	1.262**	1.098**
	(0.408)	(0.467)	(0.497)	(0.515)	(0.440)
Spill. def.	# of firms	# of firms	# of firms	# of firms	# of firms
Firm-year FE	Yes	Yes	Yes	No	No
Firm FE	No	No	No	Yes	Yes
Year FE	No	No	No	No	No
Product FE	Yes	No	Yes	No	Yes
District FE	Yes	No	No	No	No
District-year FE	No	No	No	No	Yes
HS2-district FE	No	Yes	No	No	No
Country-district FE	No	No	Yes	No	No
Product-district-year FE	No	No	No	Yes	No
Num. obs	24,579	24,579	24,579	24,579	24,579
Adjusted R2	0.402	0.394	0.451	0.383	0.220

Notes: The dependent variable is a binary indicator taking the value of 1 if the export relationship initiated by a manufacturer started to export product p to country d at time t survived beyond its k year(s) (i.e. lasted at least $k+1$ years) as defined in (9). Sample consists of all new export relationships initiated between 2002 and 2007 by Peruvian exporters of manufactured products. Spillovers at the firm-product definition are calculated as the number of firms exporting the same HS4 product unless otherwise stated. Other regressors included in the baseline specification in column (3) of table 5 are included but not reported for space considerations. Specifications that do not include firm-year fixed effects include the firm's lagged total export sales as a proxy for total factor productivity to control for firm-level heterogeneity. The estimated parameters are multiplied by 100 for ease of interpretation. Standard errors in parenthesis are clustered at the district-market level. All count regressors are in logarithms and lagged one period. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$