WPS6062

Policy Research Working Paper 6062

Together We Stand?

Agglomeration in Indian Manufacturing

Ana M. Fernandes
Gunjan Sharma

The World Bank
Development Research Group
Trade and Integration Team
May 2012

Policy Research Working Paper 6062

Abstract

This paper uses plant-level data to examine the impact of industrial and trade policy reforms on the geographic concentration of manufacturing industries in India from 1980 to 1999. First, the research shows that de-licensing and liberalization in foreign direct investment significantly reduced spatial concentration, but trade reforms had no significant effect on spatial concentration. Second, plants respond differently to policy reforms based on their size. Liberalization in foreign direct investment and de-licensing caused small plants to disperse, while trade liberalization had the opposite effect. However, for large plants trade liberalization led to lower spatial concentration.

This paper is a product of the Trade and Integration Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at afernandes@worldbank.org and sharmag@missouri.edu.

Together We Stand? Agglomeration in Indian Manufacturing

Ana M. Fernandes* and Gunjan Sharma‡

JEL classification:F10, F15, O2, R12. Keywords: Trade, Foreign direct investment, delicensing, agglomeration, spatial distribution, India

*World Bank Development Research Group, afernandes@worldbank.org

†University of Missouri, sharmag@missouri.edu

‡Support from the governments of Norway, Sweden and the United Kingdom through the Multi-Donor Trust Fund for Trade and Development is gratefully acknowledged. We would like to thank Devashish Mitra and Hiau Looi Kee as well as seminar participants at University of Missouri, the Midwest Trade Meetings, the IGC Urbanization Conference, the North American Summer Meeting of the Econometrics Society, and the World Bank. The findings expressed in this paper are those of the authors and do not necessarily represent the views of the World Bank.

1 Introduction

Recent literature has shown that agglomeration and economic density are important drivers of productivity and economic growth (e.g., Ciccone and Hall (1996), Cingano and Schivardi (2004), Lall et al. (2004), Brulhart and Sbergami (2009)). However, the clustering of production may lead to increased regional concentration of income if labor is imperfectly mobile (Hanson and Harrison (1999), Topalova (2004)). This in turn may lead to a regional concentration of poverty and increasing inter-regional disparities, particularly when labor mobility across regions is limited. Both these factors are extremely relevant for India where the spatial distribution of manufacturing has several interesting and peculiar features.[1]

Firstly, the distribution of industries across Indian states is highly skewed. In 1985, the locational gini coefficient for Indian manufacturing was 0.7 as compared to 0.25 for Chinese manufacturing.[2] Secondly, the spatial distribution of manufacturing industries in India experienced significant changes in the 1980s and 1990s. The growth rate of the various states' share of manufacturing employment exhibits differing trends not only across the two decades but also across smaller sub-periods. Thirdly, there is significant heterogeneity in agglomeration patterns across Indian industries.[3] Fourth, from 1950 to 1990 industrial policy - in particular, licensing - was explicitly used by the government to make the spatial distribution of industry "less unequal". These features gain even more significance in face of the major

[1] Regional trends in inequality in India have been a cause of concern, particularly in view of the massive economic changes the country experienced since the mid-1980s. Deaton and Dreze (2002) show evidence of divergence in per capita consumption across states and of growing urban-rural inequalities, both within and across states. See Pal and Ghosh (2007) for a detailed review of evidence of increasing regional disparities in India.

[2] For each country, the locational gini coefficient measures the inequality in a measure of regional specialization of industry j and its formula is given by $g_j \equiv Gini_j(r_{js})$ where $r_{js} \equiv \frac{\frac{L_{js}}{L_s}}{\frac{\sum_{s=1}^{M} L_{js}}{\sum_{s=1}^{M}\sum_{j=1}^{J} L_{js}}}$ with L_{js} being employment of industry j in geographic area s and L_s being total manufacturing employment in geographic area s, $s = 1, ..M$. The coefficient for China is taken from Ge (2009) while that for India is the average across the Gini coefficients for 2-digit industries calculated by the authors.

[3] Unreported figures (available from the authors) show that for example employment shares in the wool, silk and man-made fiber industry (NIC 24) grew at fairly similar rates across Indian states between 1980 and 1990 providing no evidence of increasing agglomeration. However from 1990 to 1999, most states experienced negative growth rates while a handful experienced extremely high positive growth rates of employment share of this industry. Thus the industry agglomerated from 1990 to 1999. On the other hand, it is difficult to say whether employment in the chemical and chemical products industry (NIC 30) became relatively more agglomerated during the 1990s compared to the 1980s. What happened was a change in the identity of the states with high growth rates of employment share of this industry from 1980s to the 1990s.

industrial and trade policy reforms and the general shift towards market-oriented economic policies that occurred in the 1980s and 1990s in India. Theory does not offer clear-cut predictions regarding the effect of industrial de-licensing, trade liberalization, and foreign direct investment reforms on spatial concentration of manufacturing industries, and hence the need for empirical analysis.

In this paper we ask the following questions. What are the determinants of the high level of spatial concentration in Indian manufacturing? How much of the changes in the spatial distribution of Indian manufacturing industries can be attributed to policy reforms, after conditioning on the traditional determinants on agglomeration - endowments, Marshallian-type linkages and knowledge spillovers? Does the spatial distribution of manufacturing plants of different sizes differ in response to policy changes? To answer these questions, we focus on indices of geographic concentration of manufacturing industries in India constructed following Ellison and Glaeser (1997) - henceforth EG indices - based on plant-level data from the Annual Survey of Industries (ASI) for the period 1980-99.

Our results show that industrial de-licensing has a significant negative impact on geographic concentration levels (a one standard deviation increase in the proportion of output de-licensed leads to a 9.3% decline in the EG index) and the same is true for foreign direct investment (FDI) liberalization (with a one standard deviation increase in the proportion of industry output open to FDI reducing the average EG index by 10.5%). However, our estimates also show that trade policy has no effect on the spatial concentration of Indian manufacturing. Further we find that there is significant size-based variation in the response of concentration to the policy as well as the traditional determinants of agglomeration. De-licensing and FDI liberalization cause small plants to disperse. Trade policy has no effect on the concentration of medium-size plants but a one standard deviation decline in the effective rate of protection leads to a 20% decline in the EG index of large plants and a 2.5% increase in the EG index of small plants. Our results are robust to heteroskedasticity-correcting FGLS techniques and controls for persistence in the agglomeration process as well as to a wide variety of fixed effects and specification tests.

A large literature provides evidence of a significant positive effect of agglomeration economies on firm-level as well as aggregate-level productivity. Thus, our finding that policy reforms in India raised geographic dispersion would imply that there were productivity losses (or at the very least, lower productivity gains as a result of the reforms) in Indian manufacturing. This implication of our results could be one explanation for the finding that only a quarter of the Indian growth miracle can be attributed to the policy reforms shown by Bollard et al. (2010). That is, the tendency of Indian manufacturing to disperse after trade and industrial policy reforms reduced its ability to take advantage of agglomeration economies, and hence reduced aggregate productivity growth.

Our study's contribution to the literature is three-fold. First, our index of agglomeration - the EG index - provides a detailed picture of spatial concentration at a micro level. Since it is computed at a relatively disaggregated 3-digit level of industries across Indian states it avoids problems associated with plants changing product mix (and hence, changing their industrial classification which would occur if we considered 4-digit industries), and problems associated with very small geographic units.[4] Second, our study is one of the first that explicitly controls for and measures the impact of three different policies - industrial de-licensing, trade reforms, FDI liberalization - in addition to controlling for public sector presence in Indian manufacturing. Third, unlike other studies which focus on cross-sections or very short panels, our data spans a 19 year period for approximately 170 industries. Hence, our measures of industrial and trade policy as well as the proxies for the traditional determinants of concentration cover a long and interesting time span. They allow us to present a complex and nuanced picture of the forces that affect spatial concentration. Moreover, the use of a long panel of industries allows us to control for industry and time fixed effects in our specifications. Also we are able to lag explanatory variables to account for the slow moving nature of the determinants of agglomeration. We account for serial correlation by estimating clustered standard errors. Lastly, to the best of our knowledge our study is the first to consider EG indices of agglomeration for different plant size categories.

The paper is organized as follows. In Section 2 we discuss the measurement of agglomeration and the evidence on the traditional determinants of agglomeration. In Section 3 we discuss the policy determinants of agglomeration - de-licensing, trade reforms, and FDI

[4] According to Mayer and Mayer (2004) the use of very small geographic units "can lead to an underestimation of agglomeration levels since it artificially separates clusters that sprawl across the border between units".

liberalization. In Section 4 we describe and summarize our data. Section 5 presents the estimation strategy while Section 6 discusses the results. Section 7 discusses the results using plant-size-based measures of agglomeration. Section 8 presents the results from alternate specifications. Section 9 concludes.

2 Agglomeration: Measurement and Traditional Determinants

2.1 Measurement of Agglomeration

Several indices can be used to measure the geographic concentration of an industry within a country: e.g., locational Gini coefficients, Hoover indices and Theil indices. Suppose that s_{js} is the share of geographic area s, $s = 1, ..M$, in industry j's total employment in the country that is, $s_{js} = \frac{L_{js}}{L_j}$. Further, x_s is the share of geographic area s in total manufacturing employment in the country, that is, $x_s = \frac{L_s}{L}$. The first measure we consider is a raw geographic concentration index for each industry j in a given time period which is given by $G_j = \sum_{s=1}^{M}(s_{js} - x_s)^2$. The index measures the degree to which the geographic pattern of employment in the industry departs from the geographic pattern of manufacturing employment in the country as a whole. The index takes a value of zero if industry employment is uniformly spread across geographic areas. Greater values of the index indicate relatively stronger geographic agglomeration of the industry. A potential problem with this raw geographic concentration index is that even if plant location was chosen completely at random it would by definition be larger in industries including only a small number of very large plants. Thus G_j could be large solely due to industrial structure.

Hence, the second and main measure we consider is the EG index proposed by Ellison and Glaeser (1997) which controls for the number and size distribution of plants, i.e., the impact of the industrial structure. The EG index is the most commonly used index in the agglomeration literature. Ellison and Glaeser (1997) develop a plant location choice model to analyze geographic concentration at the industry-level that captures both the random agglomeration that would be generated if plants chose their location by throwing darts on a board as well as additional agglomeration caused by localized industry-specific spillovers and natural advantages. Drawing upon their model they propose an index of geographic concentration of employment in an industry that controls for industry characteristics, namely for the size distribution of plants in the industry.[5] Indeed, one of the key advantages of the EG index is that it allows for cross-industry comparisons.

Considering the aforementioned raw geographic concentration index G_j for industry j and its plant size distribution as measured by the Herfindahl index, $H_j = \sum_{i=1}^{N} z_{ij}^2$ where z_{ij} is plant i's share in employment of industry j, the EG index for industry j in a given time period is defined by Equation 1:[6]

$$\gamma_j \equiv \frac{(\sum_{s=1}^{M}(s_{js} - x_s)^2) - (1 - \sum_{s=1}^{M} x_s^2)\sum_{i=1}^{N} z_{ij}^2}{(1 - \sum_{s=1}^{M} x_s^2)(1 - \sum_{i=1}^{N} z_{ij}^2)} \equiv \frac{G_j - (1 - \sum_{s=1}^{M} x_s^2)H_j}{(1 - \sum_{s=1}^{M} x_s^2)(1 - H_j)} \tag{1}$$

The subtraction of the term $(1 - \sum_{s=1}^{M} x_s^2)H_j$ in Equation 1 accounts for the aforementioned problem that G_j would be larger in industries with just a few large plants even if plants chose their locations randomly. That term corresponds to the level of concentration that would be expected (in the model of Ellison and Glaeser (1997)) if plants in the industry chose their locations independently and randomly and both industry-specific spillovers and natural advantages played no role. Therefore, the EG index measures the concentration of employment in an industry over and above the level of concentration that would have prevailed if plants were distributed completely at random. Greater values of the EG index are associated with greater geographic concentration of the industry.[7] Dumais et al. (2002) show that the index can be generalized to a dynamic setting. Note that we compute the Herfindahl index H_j based on our plant level data for Indian industries, rather than take it from external sources as in other studies. This is a particularly important advantage of

[5]The main prediction of the model is that the relationship between average levels of concentration and industry characteristics is the same whether concentration is caused by from spillovers, from natural advantage, or from a combination of both. This prediction implies that one can construct an index that controls for industry differences, regardless of the cause of concentration.

[6]The Herfindahl index is defined as usual covering all plants in the industry without a geographic dimension.

[7]A value of zero for the EG index indicates that industry employment is as concentrated as would be expected from a random location process. Ellison and Glaeser (1997) suggest that industries with EG index values higher than 0.05 are highly concentrated while those with EG index values below 0.02 are not very concentrated.

our index since it allows us to compute the Herfindal index at the same level of industrial disaggregation as the raw geographic concentration index G_j.

2.2 Traditional Determinants of Agglomeration

Mayer and Mayer (2004) and Combes and Overman (2004) review in detail a large number of studies that analyze the determinants of agglomeration measured by the rate of growth of employment of an industry in a particular location or by some agglomeration index for an industry. These studies focus on knowledge externalities (Glaeser et al. (1992)), labor pooling (Overman and Puga (2009), Amiti and Cameron (2007)), input networks (Holmes (1999), Rosenthal and Strange (2001)), demand networks (Hanson (1996), Hanson (1997)), natural advantage, and increasing returns to scale (Kim (1995), Amiti (1999), Haaland et al. (1999)) as factors that affect plant location. The intuition behind these various mechanisms for agglomeration is the following.

First, plants will tend to locate in areas that provide localization economies (for example, areas near raw material suppliers, areas with existing input networks, areas where the supply of the types of workers that the plant needs is plentiful, areas with good infrastructure) and urbanization economies (for example, areas where sectors and plants that are a source of demand for the plant's product are located). This means that an industry which is more reliant on, say, domestically purchased inputs, or on transportation will tend to locate in areas that have existing input networks. Thus such industry will tend to be more agglomerated than an industry that is less dependent on inputs and roads.

Second, industries with increasing returns to scale can lower their costs of production and raise productivity if they locate production in a few locations only. Hence, industries characterized by increasing returns to scale will tend to be more geographically concentrated than other industries. However Haaland et al. (1999) argue that while theory provides an unambiguous prediction that industries characterized by a higher degree of returns to scale will tend have higher levels of absolute concentration (i.e., when an industry is unequally distributed across geographic units), no such prediction holds for the case of relative concentration. That is, a higher degree of returns to scale may not affect how different the geographic spread of an industry is relative to the average spread of industries between geographic units. As Mayer and Mayer (2004) report, there is little evidence that the degree of returns to scale affects agglomeration. In fact, Haaland et al. (1999) find that higher degrees of economies of scale are associated with lower relative and absolute concentration levels.

Third, plants will also choose location based on dynamic externalities - the Marshall (1890), Arrow (1962), Romer (1984) or MAR externalities. As plants tend to locate near one another, the extent of knowledge spillovers increases, leading to greater productivity of the industry. This results in a higher rate of growth of employment of that industry and that location. Other dynamic externalities include those modeled by Porter (1991), that relate the degree of competition in a location to higher innovation and hence to greater knowledge spillovers and greater concentration, as well as others modeled by Jacobs (1969) which relate variety and diversity in a location to greater knowledge spillovers, and hence to greater concentration in that location. Our specification will include controls for all these traditional determinants of agglomeration. as described in Section 5.

3 Policy Determinants of Agglomeration

Before describing in detail the policy reforms in India, we discuss existing evidence regarding the explicit role of government policy for the geographic concentration of manufacturing. Lall and Chakravorty (2005, 2007) study the determinants of the location of new private and public sector industrial investments in India as structural reforms proceeded between 1992 and 1998. Their results show that new private sector industrial investments were biased towards existing industrial clusters and coastal districts and depended on the size of the investments in the same location prior to the reforms and on the size of the investments in neighboring locations as reform proceeded. The authors infer that structural reform led to increased spatial inequality in terms of industrialization. These studies provide some interesting and important stylized facts about the location of manufacturing industry in India. However, due to the nature of the data used they are unable to analyze long-term patterns and determinants of agglomeration as we do in this paper. Bai et al. (2004) investigate the evolution of the Hoover index of agglomeration constructed using Chinese industry-level data from 1985 to 1997. They estimate the impact of local protectionism (i.e., the fact that local governments shelter certain industries in the region from competition) on agglomeration. Their results show that industries with high past profit margins and high shares of state ownership are less geographically concentrated.[8] Lu and Tao (2009) address that same question using Chinese industry-level data from 1998 to 2005 and show that there is less geographic concentration in industries with higher shares of employment in state-owned enterprises. Fujita and Hu (2001) show an increasing income disparity between China's coastal and interior regions and attribute it to the clustering of manufacturing and the self-selection of FDI into coastal Special Economic Zones. Ge (2009) demonstrates that access to foreign trade and FDI is a driving force of unbalanced spatial distribution in China. In particular, industries dependent on foreign trade and FDI are more likely to locate in regions with better access to foreign markets.

Martincus and Sanguinetti (2009) estimate the impact of the interaction between tariffs and distance to the capital city on the regional employment shares of manufacturing industries in Argentina (for years 1974, 1985, and 1994). They find that lower tariffs were associated with the de-concentration of manufacturing activities away from Buenos Aires and its surrounding region, which they interpret as indicative that with the opening to trade, demand and cost linkages weakened and agglomeration diseconomies such as high commuting costs or high land rents prevailed.[9] Using measures of openness based on trade volumes rather than trade policies, Martincus (2010) estimates a similar econometric specification to explain the distribution of Brazilian industries across states and shows that more open industries tended to locate in states near the country's largest trading partner Argentina, and this tendency increased with the trade liberalization of the 1990s. These studies demonstrate that industrial policies can have large and long term impacts on spatial concentration and, potentially, on interregional wage and income inequality. In the sections below, we describe the policies that were in place in India prior to the reforms and whose liberalization we expect

[8]The mechanism behind their findings is that in the fiscally decentralized Chinese regime local governments have an incentive to protect industries with high profit margins and state-owned enterprises because they are able to expropriate some of their profits through ad-hoc taxes and fees.

[9]This finding is obtained in specifications that control for comparative advantage and input-output (IO) linkages factors through interactions of industry characteristics (e.g., intermediate input intensity) and region characteristics (e.g., natural resource endowments).

to have an effect on the spatial distribution of manufacturing.[10]

We consider below three types of industrial policies – licensing of private industry, barriers to international trade, and regulation of FDI – and their potential impact on the spatial distribution of manufacturing industries. Note that these policy reforms were implemented differentially across industries and time, but not differentially across geographic regions.

3.1 The License "Raj"

Since the 1950s, industrial licensing was one of the major methods to control private enterprise in India and to direct private capital into desirable industries. Under the Industries (Development and Regulation) Act of 1951, all factories that were already operating or wished to operate in a specified list of industries were required by the government to obtain a license to continue or begin production.[11] Licenses were issued only by the central government and affected several aspects of a plant's operation. Almost by definition, the licensing regime controlled entry into the industry and hence the amount of competition faced by a plant. A license also specified the amount of output that a plant could produce. Importantly, licenses were conditional on the proposed location of the project and permission was required to change locations.[12]

During the 1970s and 1980s the Indian licensing regime was used as an important instrument for determining location decisions.[13] The administration of the licensing regime allowed the central government to give positive weight to proposals intending to locate production in backward areas and to give negative weight to proposals located in/near identified metropoli-

[10]For more details, see Krishna and Mitra (1996), Sharma (2006b), Sivadasan (2009) and Topalova and Khandelwal (2011) and [illegible] and Sharma (2011).

[11]Factories are defined as (i) enterprises that do not use power but employ more than 100 workers or (ii) enterprises that use power and employ more than 50 workers.

[12]In addition, the exact nature of the item to be produced was also specified and the plant needed permission or another license to change its product mix. Even the type of technology and inputs that the plant could use in production (though not specified on the license) was determined because the most crucial raw materials (steel, cement, coal, fuel, furnace oil, railway wagon movements, licenses to import inputs) were controlled by the government and the plant needed to get annual allotments of these for its production. While deciding on a license, the considerations of the licensing committee were mainly macro-economic in nature and had little to do with the project's merits. For example, policy-makers disdained variety and thought that competition was wasteful.

[13]Appendix A provides more details on how the industrial policy regime (in particular, licensing) was used to impact the spatial distribution of manufacturing industries.

tan and urban areas.[14] Marathe (1989) finds that these policies were quite effective: e.g., in 1975 and 1976 45% of approved licenses corresponded to projects located in backward areas but at the same time all the other licenses went to only four industrialized states. Thus the industrial distribution in India was tending towards bi-modal. In the 1980s, the Indian government started to relax the licensing regime by de-licensing certain industries. Table 1 shows the percentage of manufacturing employment that was de-licensed in selected years in the 1980s and 1990s as well as the cumulative percentage of total employment, output and capital that was de-licensed. This piecemeal approach to reforming industrial policy continued throughout the 1980s. In 1991, the Indian economy faced a balance of payment (BOP) crisis and was forced to take loans from the International Monetary Fund (IMF). Under pressure from the IMF, the most significant de-licensing episode occurred — industrial licensing was removed for 84% of manufacturing output.

Some aspects of the first phase of reforms in the 1980s are noteworthy. First, the impact of de-licensing in a particular industry was different for differently sized plants. The licensing regime included a size-based exemption which meant that plants below a certain size threshold were exempt from the licensing provisions (Sharma (2006a,b)). However these small plants still faced locational constraints - for example, they could not locate near any large Indian city - but these constraints were milder than those imposed on large plants. Thus, within a given industry, one can expect different patterns of spatial concentration for plants of different sizes. Second, de-licensing of industries in the 1980s was accompanied by continued restrictions on plant location. In particular, large plants could avail of the de-licensing only if they located in a set of "backward areas" that is, those with little or no manufacturing industry. This again points to the possibility of a size-based response of spatial concentration to de-licensing of an industry.

The potential impact of de-licensing on concentration in India is ambiguous. On the one hand, the industrial policy regime was willfully used to settle industry in backward, non-industrialized areas. Hence, after an industry was de-licensed, plants would tend to choose locations based on economic criteria: e.g., areas with good infrastructure, availability of

[14] The administration of the licensing regime was characterized by trials by a licensing committee, a high degree of centralization, and a tradition of placing microeconomic constraints on firms regarding output, technology and input use.

input networks, and an appropriately skilled labor force. This would suggest that industrial de-licensing should be associated with an increase in agglomeration. On the other hand, industrial policy was creating artificial clusters of industry in backward areas. As licensing requirements were removed, these clusters would likely break up, leading to a decline in agglomeration levels.[15] Another issue to note is that de-licensing in the 1980s was conditional on plant size and location. That is, a large plant was considered de-licensed only if it located in certain "backward" areas. Thus, de-licensing might have led to even more geographic dispersion in plant location in India.

3.2 The Tariff-Quota "Raj"

Prior to reform India's trade regime was one of the most restrictive in Asia. The regime consisted of high nominal tariffs, a complex import licensing system, an actual user policy that restricted imports by intermediaries, restrictions of certain exports and imports to the public sector, phased manufacturing programs that mandated progressive import substitution, and government purchase preferences for domestic producers (Topalova and Khandelwal (2011)). After the BOP crisis of 1991, there were major changes in both the tariff and non-tariff barrier levels applied to Indian industries, as well as in the methods by which the trade regime was implemented. Non-tariff barriers were rationalized and scaled down: e.g., 26 import licensing lists were removed and a single list that contained prohibited imports (the 'negative' list) was established. Average tariffs fell by 43 percentage points between 1990 and 1996, and the standard deviation of tariffs fell by 50%. The Rupee was devalued 20% against the dollar in 1991 and in 1993 India adopted a flexible exchange rate regime. The rationalization of tariff and non-tariff barriers continued into the early 2000s.

The implications of trade liberalization for the regional distribution of industries within a country have been examined by several New Economic Geography (NEG) models, where the location choices of firms and consumers are determined by opposing centripetal and centrifugal forces.[16] The agglomeration (centripetal) forces result from the interaction between

[15]Ideally, we would like to consider, in addition to the licensing regime, the direct impact of subsidy and concessional finance schemes that were offered at the state level but data for these are not available.

[16]For excellent reviews of the literature on the spatial effects of international trade see Mayer and Mayer (2004) and Brülhart (2010). The essential features of NEG models are: (1) increasing returns to scale

IRS, market size (location of demand), the costs of trading across regions and countries, and/or backward and forward linkages among producers. The dispersion (centrifugal) forces differ across NEG models, in some cases they arise from the costs that firms face to reach an exogenously dispersed demand of immobile consumers (working in the agricultural sector) while in other cases they arise from congestion costs (rent and commuting costs) associated with large industrial agglomerations.

Several NEG models generate the prediction that trade liberalization favors the geographic concentration of manufacturing activities within the liberalizing country. Paluzie (2001) and Monfort and Nicolini (2000) derive this prediction by extending the models by Krugman (1991) and Krugman and Venables (1995) to the set-up of, respectively, a country with two symmetric regions liberalizing its trade with the rest of the world and a country with two symmetric regions liberalizing trade against a foreign country (also with two symmetric regions). As international trade is liberalized, (i) access to foreign demand (exports) lowers the incentives for domestic firms to locate near domestic consumers that now represent a smaller share of sales and (ii) the presence of foreign supply (imports) lowers the incentives for domestic firms to locate near other domestic firms for input-output (IO) linkages since foreign firms now represent a larger share of supply to domestic consumers. The incentive for a firm to locate away from domestic competitors (in the periphery) is provided by the possibility of being sheltered when serving the local market. With trade liberalization, competition in the periphery comes also from foreign firms and makes the dispersion of manufacturing activities less attractive. Brülhart et al. (2004) and Crozet and Koenig (2004) consider regions that differ in their distance to the rest of the world: a border region and an interior region. The main prediction from their NEG models is that trade liberalization fosters spatial concentration of manufacturing in the border region with better access to international markets. The novel forces in these models are that domestic firms may be attracted to the border region to reap the full benefits from improved access to foreign demand and domestic

(IRS) internal to the firm (e.g., due to indivisible fixed costs) and monopolistic competition generally of the Dixit and Stiglitz (1977) type in the manufacturing sector; (2) costs of trading outputs produced and inputs used by firms across distances; (3) endogenous firm location decisions ; (4) endogenous location of demand. Despite a vast NEG literature following Krugman (1991), due to technical difficulties in characterizing the equilibrium distribution of economic activity, only recent models that assume many regions within a country and distinguish across regions and countries allow researchers to characterize the response of agglomeration to international trade liberalization.

consumers face an incentive to agglomerate in the border region to access imported goods.

However, another set of NEG models generate the prediction that trade liberalization favors the geographic dispersion of manufacturing activities within the liberalizing country. Krugman and Elizondo (1996) consider a model with symmetric regions and congestion costs and show that upon opening to trade, the importance of foreign demand and foreign supply increases and the weight of backward and forward linkages decreases. Hence, the location of domestic firms and consumers in a larger domestic market becomes less important.[17] Since congestion costs are present and independent of trade openness, their strength under a large set of parameter values leads to the dispersion of manufacturing activities. Behrens et al. (2007) obtain the same prediction in a model with endogenous competition effects.[18] In this model the agglomeration of competing firms reduces their market power, which imposes downward pressure on their local markups and this pro-competitive effect acts as an additional dispersion force. Summing up, the sign of the relationship between trade liberalization and agglomeration of manufacturing firms within a country is theoretically ambiguous. Hence, empirical evidence is necessary to determine that sign in the case of India.

3.3 FDI Liberalization

The industrial policy regime in India from the 1970s onwards controlled foreign direct investment. Prior to 1991 foreign ownership rates were restricted to be below 40% in most industries (Sivadasan (2009)). In addition, restrictions were placed on the use of foreign brand names, on remittances of dividends abroad, and on the proportion of local content in output. Liberalization of FDI occurred only after the BOP crisis of 1991. Foreign ownership of up to 51% was allowed for a group of industries and other restrictions on brands, remittances and local content were relaxed.

To our knowledge, the theoretical effects of FDI liberalization on the agglomeration of

[17] The authors use their model to rationalize the existence of giant cities in developing countries as a by-product of import substitution policies. The prediction is obtained assuming that inter-regional transport costs within the country are lower than international transport costs and the autarkic equilibrium is characterized by a spatial concentration of economic activities due to backward and forward linkages.

[18] The paper extends to a spatial setting the monopolistic competition model of Ottaviano et al. (2002) where firms face a variable demand elasticity. In this model, the exogenously dispersed demand of unskilled workers is the dispersion force and trade liberalization has a pro-competitive effect i.e., firm mark-ups fall with the number of local producers rather than being fixed as in most other NEG models.

manufacturing industries have not been explicitly studied, but some insights from the NEG literature can be borrowed to conjecture about those effects. FDI liberalization in an industry is expected to bring the entry of new foreign firms, with a direct effect on the concentration or dispersion of manufacturing across regions, but can also affect the location decisions of (new or incumbent) domestic firms, with an indirect effect on the concentration of manufacturing across regions.

Focusing first on the entry of new foreign firms, NEG models suggest that if this new FDI enters the country to supply the domestic market, it would tend to locate close to large markets to economize on inter-regional transport costs. If those transport costs are low, however, then foreign firms would be indifferent about their location since they could serve domestic customers from anywhere, possibly contributing to the geographic dispersion of the industry. The theoretical model and the evidence in Amiti and Javorcik (2008) show that market access was crucial in determining the location of new FDI inflows across Chinese provinces. For new FDI aimed at exploiting production advantages such as cheap unskilled labor and targeting export markets, the optimal location choice is not clear. On the one hand, a new foreign firm may have an incentive to locate away from main industrial centers where wages are higher - leading to industrial dispersion - but on the other hand it could prefer to locate near major centers to use a pool of skilled labor or to be near ports or major borders - increasing industrial agglomeration. Amiti and Javorcik (2008) show that lower production costs also played an important role in attracting new FDI inflows to certain Chinese provinces. For new FDI of both types, the degree of usage of domestic versus imported inputs could also affect location choices. A new foreign firm relying heavily on domestic IO linkages would have an interest in locating close to suppliers which may or may not be already located in highly concentrated industrial areas. Amiti and Javorcik (2008) show that supplier access was critical in determining the location choices of new FDI inflows in China.[19]

Focusing next on the location decisions of domestic firms in response to the entry of foreign firms, on the one hand domestic firms would have an incentive to locate away from foreign firms if those are competitors in serving the domestic market as in NEG models. But on the other hand, the possibility of IO linkages with (and expected knowledge spillovers

[19]See Ottaviano and Thisse (2008) for a survey of the rationales for location of new FDI.

from) foreign firms could lead domestic firms to locate near foreign firms. Since the location of foreign firms themselves near or far from existing agglomerations is not clear-cut, the effects for domestic firms' choices of location are also not clear-cut. In policy-making circles it is often believed that FDI can stimulate the geographical concentration of activities for example into agglomerations of small firms acting as suppliers to multinationals (Propris and Driffield (2006)). In some cases, new foreign firms are explicitly attracted to be part of such clusters as is the case for Special Economic Zones in China. But at the same time many policy-makers believe that FDI can support the regeneration of less-favored regions, which would result in a dispersion of industrial activity. The large subsidies provided to FDI firms to locate in laggard regions in developed and developing countries alike illustrate this perspective (Haskel et al. (2007)).

As in the case of trade liberalization, the overall effect of FDI liberalization on industrial agglomeration is ambiguous depending, among others, on the motive for FDI, the degree of local sourcing of inputs by new foreign firms, and the pre-existing geographical distribution of industries.

3.4 Public Sector Reservation

The mixed economy framework prevailing in India prior to reforms mandated a large and growing role for the public sector. Certain important industries were reserved exclusively to the public sector. Additionally, the location of public sector enterprises was used to further the goal of industrializing the backward regions of the country, disregarding economic conditions. It is therefore possible that the spatial distribution of Indian industries with a large share of public sector enterprises differs systematically from that of comparable industries with less public sector presence. Moreover, the economic reforms of 1991 brought a paradigm shift in political and public views regarding the role of the public and the private sector in India. As Table B.1 shows, the average share of output produced by state-owned enterprises declined from an average of 30% in the 1980s to 15% in 1994, and then to 13% in 1999.

4 Data

We use plant-level data from the Annual Survey of Industries (ASI) conducted by the Central Statistical Organization (CSO), a department of the Ministry of Programme Planning and Implementation of the Government of India to compute our agglomeration indices for all consecutive years in the period 1980-81 to 1999-00, with the exception of 1995-96 when the survey was not conducted. The length of our data allows us to cover all the Indian reforms of the 1980s as well as the major reform episode of 1991. The survey covers all factories registered under the Factories Act of 1948. The ASI frame can be classified into 2 sectors: the 'census sector' and the 'sample sector'. Units employing more than 100 workers constitute the census sector. Units in the census sector are covered with a sampling probability of one while units in the sample sector are covered with sampling probabilities lower than one (one-half until 1987-88 and one-third after 1987-88).[20]

As Bollard et al. (2010) explain in detail in their appendix, there are substantial caveats to the use of these data. Sampling schedule changes in 1997 led to drastic changes in sample size: the number of total plants covered and the number of plants in the census sector dropped sharply.[21] There is a sharp increase in the standard deviation of average employment in both the census and the sample sectors, causing us to worry about heteroskedasticity. In order to correct for the noise and heteroskedasticity, we will weight all descriptive statistics and regressions by 3-digit industry-year total employment. That is, we will give more weight to larger, presumably less noisy industries.[22] As an alternative to employment weights, we will also use FGLS techniques to weight each observation by the inverse of its variance relative to industry and year means. That is, we will downweight observations which are very far from industry and year means. We discuss the technique in detail below. In order to measure changes in industrial policy, we use the detailed dataset of industrial policy in India constructed by Sharma (2006a,b) and used in Chamarbagwalla and Sharma (2011). This dataset identifies which 4-digit industries underwent reform in terms of freedom from licensing requirements in each year from 1970 to 1990. Table 1 shows the quantum of de-

[20]Note that the ASI data is a series of repeated cross-sections of plants, not a panel of plants, thus it does not allow one to identify entry and exit.

[21]See Appendix Figures B.1 and B.2.

[22]Appendix Table B.1 shows the weighted averages of the key variables used in our analysis.

licensing that took place during the sample years and brings forward two important points. The first is that the reforms of the 1980s were quite significant in terms of the percentage of manufacturing output, employment and capital affected. Cumulatively, 25% of output and employment had been de-licensed as of 1990. Hence, studies that ignore the pre-1991 changes in the licensing regime provide misleading estimates of the impact of the 1991 reforms. The second is that de-licensing in 1991 was not across the board as is the common assumption in most studies. After 1991, 1[illegible]% of manufacturing output and 11% of employment remained under compulsory licensing though some of these industries were gradually de-licensed in 1993 and 1994. The de-licensing measure used by Sharma (2006a) is a dummy variable at the 4-digit industry level denoted by De_{kt} that is equal to one in all years greater than equal to year t if industry k was de-licensed in year t. Since our EG index is computed at the 3-digit industry level, we need to calculate the proportion of employment in the 3-digit industry that was de-licensed as of year t. Hence, our measure of industrial de-licensing is given by $DEL_{jt} = \frac{\sum_{k=1}^{N_j} De_{kt} Y_{kt}}{\sum_{k=1}^{N_j} Y_{kt}}$ where k indexes the 4-digit industries within the 3-digit industry j and Y refers to employment.

Data on trade policy are obtained from Das (2003). The author computes the effective rate of protection (ERP) for 3-digit Indian manufacturing industries in four sub-periods: 1980-81 to 1985-86, 1986-87 to 1990-91, 1991-92 to 1994-95, and 1995-96 to 1999-00.[23] We also use tariff data from Topalova and Khandelwal (2011) which are available for every year but only from 1987 to 1999 therefore reducing our sample size. However, these data have rich cross-sectional and time-series variation - which the ERP data lack. We estimate all our main specifications with both measures of trade protection. We use the FDI deregulation variable from Sivadasan (2009) which is a dummy variable indicating which 4-digit industries were FDI-deregulated in 1992.[24] We combine this dummy variable with the ASI data to calculate the proportion of employment in a 3-digit industry that was exposed to FDI. Hence, our measure of FDI deregulation is given by $FDI_{jt} = \frac{\sum_{k=1}^{N_j} FDI_{kt} Y_{kt}}{\sum_{k=1}^{N_j} Y_{kt}}$ where k indexes the 4-digit industries within the 3-digit industry j and Y refers to employment.

[23] Das (2003) calculates these ERP measures of trade for 72 3-digit industries. In order to use all industries in our analysis, we use the average of his ERP measure for the corresponding 2-digit industry and in some cases, the economy-wide average. Note that for any given industry, the ERP has a constant value in all years within each of the four sub-periods.

[24] The announcement of this reform was made in August 1991, but its implementation began only in 1992.

5 Estimation Strategy

In order to calculate the raw agglomeration index G for each 3-digit industry and year we aggregate the plant-level employment data to the 3-digit industry and state level in each year weighting by sampling weights. To obtain the EG index for each 3-digit industry and year using Equation 1, we make use also of a Herfindahl index for each industry and year calculated using plant-level data on market shares.[25] Table 2 shows that the mean of the EG index rose from 1980 to 1990 and then declined in 1999. Thus Indian manufacturing agglomerated during the 1980s and dispersed during the 1990s. Appendix Tables B.2 and B.3 present the ten most and least agglomerated industries in 1980 and the evolution of their EG indices over the next 19 years.

Our main specification modifies that used by Rosenthal and Strange (2001) based on two important considerations. First, the spatial distribution of manufacturing industry is a slow moving process. This is bound to be the case although both the raw agglomeration index G as well as the EG index are expected to respond fairly quickly to changes in their determinants. That is because these indices change not only due to entry/exit of plants, but also due to changes in employment by existing plants. However, it does take time for existing plants to change location or for new plants to emerge in response to changes in policies or industry characteristics. Hence, we allow all our independent variables to affect agglomeration with a lag. Second, spatial concentration is likely to be persistent: industries that were more concentrated in the past may tend to be systematically more concentrated in the present and in the future. We account for a very general form of serial correlation in concentration by clustering the standard errors at the 3-digit industry level. The main specification we

[25]The ASI data do not allow us to identify geographic units smaller than states: the data includes a numeric district variable but no label to go with it. Hence it is not possible to identify which district corresponds to a particular value of the district variable. However, we were able to identify districts using matching techniques to create a pseudo-panel of plants and presuming that a given plant does not change location over time. Once districts can be identified and followed over time we are able to calculate the EG index using the district as the geographic unit. For most industries we find that the values of the EG index calculated at the district level tend to be smaller compared to those of the EG index calculated at the state level. This is consistent with the findings in other studies (Head and Mayer (2004)). These results can be obtained from the authors upon request.

estimate is the following where j denotes a 3-digit industry and t denotes a year:

$$\begin{aligned} G_{jt} or EG_{jt} &= \beta_1 DEL_{jt-k} + \beta_2 ERP_{jt-k} + \beta_3 FDI_{jt-k} + \beta_4 PUB_{jt-k} \\ &+ \beta_5 MATS_{jt-k} + \beta_6 INVEN_{jt-k} + \beta_7 LPOOL_{jt-k} + \beta_8 IRS_{jt-k} \\ &+ \beta_0 + \alpha_{j'} + \delta_t + \epsilon_{jt} \end{aligned} \quad (2)$$

The main coefficients of interest in Equation 2 are those on the policy variables: industrial de-licensing or deregulation (DEL), trade protection measured by the log effective rate of protection (ERP) or by log tariffs (TAR) in some specifications, and FDI liberalization (FDI). Our main specification also controls for traditional determinants of the geographic concentration of an industry related to the importance of localization economies. Following Rosenthal and Strange (2001) we proxy externalities from labor pooling ($LPOOL$) by the wage bill share of skilled (non-production) workers in the industry. Industries benefit from positive externalities via shared input-output networks, proxied by the log of materials per shipment ($MATS$). We use the log of inventories of finished goods per shipment to proxy for an industry's dependence on transport networks ($INVEN$) and average real capital per plant in an industry to proxy for the returns to scale in the industry (IRS). Capital is measured as the book value of average fixed capital owned by the plant. Summary statistics and definitions of these determinants of agglomeration are provided in Appendix Table B.1. As Rosenthal and Strange (2001) point out, the coefficients on these industry characteristics reflect the equilibrium relationship between geographic concentration and localization and urbanization economies. Thus, $LPOOL$, $MATS$ and $INVEN$ affect geographic concentration but are also affected by it. Hence, the coefficients on these variables cannot be interpreted as causal, though the problem may be ameliorated by the use of lagged values of these variables. In either direction of causality, the relationship hinges on cost reduction: geographic concentration reduces the costs of labor and inputs and because of this, industries sensitive to labor and input costs tend to concentrate. Thus, there is still valuable information to be gleaned from the coefficients on the variables capturing localization economies.

As most other studies on the determinants of agglomeration, we do not have appropriate measures to capture two other potentially important determinants: natural advantage and knowledge spillovers. To address the potential omitted variable bias problem that may result, we control for industry fixed effects at the 2-digit industry level in our specifications ($\alpha_{j'}$).[26] We also include year fixed effects (δ_t) in all our specifications to account for secular changes in the geographic concentration of manufacturing industries in India. In particular, these year fixed effects may capture generalized improvements in infrastructure that led to more possibilities for geographic concentration. To reduce the influence of outliers and to control for heteroskedasticity, we weight each industry-year observation by its size (measured as total employment) and inversely by its estimated variance relative to the industry-year mean. In the case of the EG index the weights are given by:

$$w3_{jt} = \frac{L_{jt}}{(E[Var_t(EG_{jt} - \hat{EG_t})]/N_t)(E[Var_j(EG_{jt} - \hat{EG_j})]/N_j)}$$

where $\hat{EG_t}$ is the average EG in year t and $\hat{EG_j}$ is the average EG in industry j.[27] To apply the GLS technique, we first estimate Equation 2 in an ancillary regression where each observation is weighted by its size and obtain the corresponding residuals. Then we project the square of those residuals on industry-year fixed effects. The predicted values from this second regression are the variance estimates used to obtain the FGLS estimator. Finally, we estimate Equation 2 again but now weighing each observation by $w3$ that is its size divided by the variance estimate. The use of $w3$ allows us to downweight both noisy industries and noisy years. In robustness specifications we will consider alternative weights that downweight just noisy industries or just noisy years.

6 Results

6.1 Main Results

The first column of Table 3 shows the results of estimating Equation 2 using the raw geographic concentration index as the dependent variable, including the one year lag of all

[26] We chose not to include fixed effects at the 3-digit level since these would soak up too much of the potentially meaningful variation in the data. Though we have a long panel of 3-digit industries, the determinants of agglomeration are industry characteristics that change slowly over time. Thus, cross-sectional variation is very important in identifying the coefficients on those determinants.

[27] In the case of the G index the weights are defined similarly replacing EG by G.

independent variables. Both de-licensing and FDI liberalization lead to a significant decline in spatial concentration of Indian industries while trade liberalization has no significant effect. The estimates using the EG index as dependent variable are presented in Column 2 and show that geographic concentration declines significantly with de-licensing and FDI liberalization conditional on standard Marshallian externalities. However, agglomeration does not respond significantly to changes in trade policy. Hence plant location in deregulated and FDI-liberalized industries is increasingly dispersed. Note that the determinants considered in our specifications explain a large fraction (57%) of the variation in spatial concentration of manufacturing in India.

In both specifications, the coefficient on the degree of returns to scale is negative and significant. This result may seem puzzling at first. But as Haaland et al. (1999) point out, the theoretical relation between returns to scale and relative concentration is ambiguous. That is, the degree of returns to scale characterizing an industry may affect where the industry locates, in the center or the periphery. Returns to scale by themselves do not allow us to draw any conclusions about how concentrated an industry is relative to other industries. One concern regarding the effects of the degree of returns to scale is that the proxy for IRS is highly correlated with the proxy for labor pooling, as shown by Table B.4. To address this potential multicollinearity problem we estimate Equation 2 excluding the proxy for returns to scale and present the results in Column 3 of Table 3. Note that we focus in what follows on specifications that use the EG index as dependent variable given its advantages relative to the G index. While in Column 3 the magnitudes of the other coefficients are similar to those in Column 2, the standard errors on some of the variables change suggesting the importance of multicollinearity. Another explanation for the negative coefficient on returns to scale relates to the actual proxy for returns to scale used. Haaland et al. (1999) aver that capital per plant is a better measure of the extent of unexploited economies of scale than of realized scale economies in an economy. Consider two industries with identical cost functions. In the industry with higher output per plant, scale economies are mostly realized and hence the marginal cost improvements from a further increase in scale are likely to be small. In the industry with lower output per plant, there is scope to monetize unexploited scale economies. Thus, one interpretation of the result is that the greater the unexploited scale economies, the less geographically concentrated will be an industry. Given our concern about multicollinearity, the results that we will present hereafter correspond to specifications that exclude the measure of returns to scale. But we should note that all results would be qualitatively similar if the measure of returns to scale was included.[28]

To subject the large effects of de-licensing and FDI liberalization in Columns 1-3 to a more stringent specification, we present in Column 4 of Table 3 the results from a specification where we replace individual 2-digit industry and year fixed effects with 2-digit industry-year interaction fixed effects. These interaction fixed effects control for time-varying industry-specific omitted variables such as technological changes or political economy factors. The estimates show that our results are robust to the control for these interaction fixed effects. [29]

In order to provide a clear economic magnitude for the coefficient estimates, we compute the elasticity of the EG index with respect to each of the independent variables using the coefficient estimates from Column 4 of Table 3 - our preferred specification - and the values of the average EG index in the first sample year 1980 (0.084). A one standard deviation rise in the proportion de-licensed output (0.49) reduces the EG index by 9.3% while a one

[28]These results are available upon request.

[29]One could argue that the estimated policy effects indicate that industries that are more geographically concentrated are more organized, have greater lobbying power, and thus were able to lobby the Indian government to not de-license them or to not liberalize FDI. That is, there might be reverse causality between geographic concentration and de-licensing or FDI liberalization. However, two points need to be made with respect to such potential reverse causality. First, there is no evidence that Indian industrialists were lobbying for particular kinds of industrial policy during the initial reforms of the 1980s. Since the reforms were not announced nor discussed within the government and legislature, it is not clear whether industrialists knew that such reforms were in the works. Further, the industrial de-licensing that took place in 1991 was a result of a BOP crisis for which political economy factors played no role. Second, even if political economy factors were important, it is not clear what type of policy stance lobbyists from geographically concentrated industries would be asking for. One the one hand, the licensing regime was a source of rents for large incumbents that would want to preserve the status quo. On the other hand, these large incumbents had to suffer the onerous conditions imposed by the licensing regime, as they could not change production levels, product mix, technology of production, nor location of production without permission. For example, Sharma (2006b) documents massive shortages of commodities like cement, scooters, and cars in India during the 1970s and 1980s, but incumbents could not respond to these profit-making opportunities due to the licensing restrictions they faced. This means that industrialists might prefer to be de-licensed. Overall, it is not clear whether geographically concentrated industries would have a unified lobbying effort in order to influence industrial policy. Assuming that political economy factors existed, they might be captured by industry fixed effects. But if they varied over time they could still bias our results: e.g., geographically concentrated industries might realize the benefits of deregulation from observing other deregulated industries within their cluster and change their lobbying efforts from maintaining the status quo to asking for deregulation. The specification in Column 4 address these concerns by including the interaction fixed effects.

standard deviation rise in FDI-liberalized output (0.2) leads to a 10.5% fall in the EG index around its mean. Interestingly, evidence from other countries (particularly China) shows that FDI deregulation leads to increased average geographic concentration. Fujita and Hu (2001) examine increasing regional (coast versus interior) disparity in growth in China particularly in the context of the location of establishments in Special Economic Zones and attribute it to the clustering of manufacturing, and the self-selection of FDI into coastal SEZs. Ge (2006) also finds evidence that foreign investments are more likely to locate in regions with better access to foreign markets. Our results show that FDI liberalization leads to increased industrial dispersion. This might highlight differences between FDI that is geared towards exports of finished products (the case in China) versus FDI that is meant to produce for the dispersed domestic (host) market which may be the case for India.

6.2 Robustness Tests

Given the ambiguity of theory regarding the effect of trade liberalization on intra-national concentration, the insignificant coefficient on ERP is not surprising. To check the robustness of this non-result, Column 1 of Table 4 shows the results from estimating our main specifications including nominal tariffs instead of ERP. The results show an insignificant effect of tariffs on agglomeration. Further, in unreported results we combine these nominal tariffs with coefficients based on input-output tables to create a proxy for input tariffs faced by each industry and find no significant effects of input tariffs on agglomeration.[30]

One potential reason for these findings could be multicollinearity between trade and FDI policy. Both reforms occurred simultaneously in August 1991 and were broad in scope. Thus in Column 2 we present a specification that includes de-licensing and FDI liberalization as the policy variables but excludes trade policy while in Column 3 we present a specification that includes de-licensing and trade liberalization as the policy variables but excludes FDI liberalization. We find that the impact of de-licensing and FDI on geographic concentration rises marginally when trade policy is excluded, but overall the results are qualitatively similar to those in Table 3. Similarly, the effects of de-licensing and the insignificant effect of trade

[30]The specification considered includes both final goods tariffs as well as input tariffs. These results are available upon request.

liberalization are maintained when FDI liberalization is dropped.

Although we control for public sector presence in all specifications so far and find the corresponding effects to be generally insignificant, it is possible that industries reserved for the public sector exhibit patterns of spatial concentration that are systematically different from those of other industries. Hence, in Column 4 of Table 4 we estimate Equation 2 including ERP again and focusing on the set of industries which were not reserved for the public sector and find that the basic qualitative results from Table 3 continue to hold.

Appendix Table B.5 presents the results from additional robustness checks. Column 1 shows the estimates weighing each industry-year observation according to its employment size which are qualitatively similar (though weaker) than those in Table 3. However, Breusch-Pagan and White tests for heteroskedasticity imply a rejection of the hypothesis of constant variance. Columns 2 and 3 show the estimates downweighting, respectively, just noisy industries and just noisy years. These corrections for heteroskedasticity reduce the standard errors for almost all coefficients when compared to those in Column 1 obtained when observations are weighted only by industry employment size. Moreover, it is important to note that all results - estimated using either weighted or unweighted regressions - are robust to restricting the sample to the years with less noisy ASI data - 1980-94.[31] Appendix Table B.5 also presents the results from estimating Equation 2 including two or three year lags of the independent variables. Allowing for this longer response of geographic concentration to its determinants does not change the coefficients much relative to those in Table 3.[32] It is interesting to note that Tables 3 and 4 show that the EG index responds reasonably quickly (within one year) to changes in policy variables. This result would be less plausible if we were examining agglomeration using the more direct and interesting approach of modeling a plant's location decision as in Mayer et al. (2010). Unfortunately, our data do not allow us to consider a plant's location decisions since the data are repeated cross-sections of plants rather than a panel. The EG index which we use essentially measures the change in a location's employment share relative to the employment share in the average location. Thus, the EG

[31]These results are available upon request.

[32]An alternative model for the evolution process of spatial concentration would include the various lags of the independent variables in a single specification. However since tests reveal a very high degree of correlation between the lags, it is not clear that the results from such a specification could be relied upon.

index and its components can increase fast in response to changes in location determinants since an employment share may increase due to a capacity expansion of incumbent plants rather than the construction of new factories. Our use of a one year lag of the independent variables does not imply any unreasonable assumptions about the birth of new plants from one year to the next.

7 Agglomeration Patterns for Plants of Different Sizes

In this section, we consider the degree of spatial concentration for different plant sizes within each industry. While industrial policy in India was geared towards macroeconomic goals, it was implemented differentially across plants, based on their size. In particular, Indian plants with a book value of fixed capital below a certain threshold (call it K_t^{large}) were exempt from licensing provisions, they did not need to take permission to enter or produce and were subject to less strict location provisions. Another category of plants with fixed capital below an even smaller threshold (K_t^{small}), entitled "small scale", were exempt from licensing in addition to having certain products reserved for their production and being subject to very lenient location provisions.

Another motivation to consider size-based EG indices is that during the 1980s, even de-licensing was administered differentially for the large plants: plants whose fixed capital was greater than K_t^{large} were de-licensed only if they located in certain backward areas. Overall it might be insightful to assess whether all other determinants considered in Section 6 affect agglomeration differentially across various plant size categories. To compute size-based EG indices, we define three dummy variables at the plant level - $S1_{it}$ equal to one if $K_{it} <= K_t^{small}$; $S2_{it}$ equal to one if $K_t^{small} < K_{it} < K_t^{large}$ and $S3_{it}$ equal to one if $K_{it} >= K_t^{large}$. The first category $S1$ is small plants, the second category $S2$ is medium-sized plants (exempt from licensing provisions but not small enough to be small scale), and the third category $S3$ is not exempt large plants. We compute the EG index for each size category in each industry, using data from the corresponding set of plants. Figure 1 presents the evolution of the average EG indices over the sample period for the three plant size categories. The figure shows significant variation in the patterns and levels of spatial concentration: small plants are more concentrated than medium-sized plants, which in turn are more concentrated than large plants. The EG index for medium-sized and for large plants also exhibits more variability relative to the EG index for small plants. All three indices show an increasing trend up to 1994, and a declining trend thereafter. But for the years 1996-99, this could be the result of noisier data rather than actual changes in spatial concentration. Table 5 presents the results of estimating Equation 3 below for each size-based EG index separately:

$$\begin{aligned} EG^s_{jt} &= \beta_1 DEL_{jt-1} + \beta_2 ERP_{jt-1} + \beta_3 FDI_{jt-1} + \beta_4 PUB_{jt-1} + \beta_5 MATS^s_{jt-1} \\ &+ \beta_6 INVN^s_{jt-1} + \beta_7 LPOOL^s_{jt-1} \\ &+ \beta_0 + \alpha_{j\prime} + \delta_t + \epsilon_{jt} \text{ for s = small, medium, large} \end{aligned} \tag{3}$$

where each of the Marshallian externalities' variables are calculated separately for the three plant size categories within each industry instead of being calculated for the average plant in each industry as in Equation 3. For example, the variable $MATS^s_{jt}$ which proxies for IO linkages is obtained as the average ratio of materials to sales for plants of size s in industry j. The same reasoning applies to the variables $INVN^s_{jt-1}$ and $LPOOL^s_{jt-1}$.[33]

Columns 1-3 of Table 5 show the results for small, medium-sized and large plants, respectively. The spatial concentration of the three types of plants responds differentially to policies. As a result of FDI liberalization in an industry, small plants disperse while medium-sized and large plants exhibit no significant response. Specifically, a one standard deviation rise in the proportion of FDI-deregulated output is associated with a 9.2% decline in the EG index for small plants. Further, a one standard deviation rise in the proportion of delicensed output is associated with a 11.4% decline in agglomeration of small plants and a 26% increase in agglomeration of medium-sized plants. Small plants tend to disperse geographically with trade liberalization: a one standard deviation fall in log ERP results in a 2.5% decline in the EG index for small plants.[34] As further evidence of differential size-based responses, we find

[33]This flexible functional form to assess size-based effects of the determinants of agglomeration is informed by recent literature that documents tremendous heterogeneity across plants even within narrowly defined industries (e.g., Foster et al. (2001)).

[34]We use the mean EG index of -0.015 for large plants and of 0.104 for small plants, as well as the 0.49 standard deviation of DEL and 0.73 as the standard deviation of ERP to calculate these economic

that large plants are significantly more spatially concentrated as a result of trade reforms: a decline in log ERP by one standard deviation reduces the EG index for large plants by 20% around its mean. However, large plants respond neither to de-licensing nor to FDI liberalization.

We also find some evidence of differential size-based effects for the traditional determinants of agglomeration. Greater labor pooling has an insignificant effect on the concentration of medium-sized and large plants but leads to more spatial dispersion of small plants. Finally, stronger public sector presence in an industry has a significant positive impact on the concentration of large plants, which presumably include some of the large, public sector enterprises. In order to facilitate hypothesis testing about the differential determinants of agglomeration across plant sizes, we pool all the size-based EG indices together and estimate a fully interacted model - defining $D2_{jt}$ to be the dummy for the medium-sized plants and $D3_{jt}$ the dummy for the large plants - given by:

$$\begin{aligned} EG^s_{jt} &= \beta_1 DEL_{jt-1} + \beta_2 ERP_{jt-1} + \beta_3 FDI_{jt-1} + \beta_4 PUB_{jt-1} \\ &+ \beta_5 MATS^s_{jt-1} + \beta_6 INVN^s_{jt-1} + \beta_7 LPOOL^s_{jt-1} \\ &+ \sum_{s=2}^{3} \gamma_{1s} DEL_{jt-1} \cdot Ds_{jt} + \sum_{s=2}^{3} \gamma_{2s} ERP_{jt-1} \cdot Ds_{jt} + \sum_{s=2}^{3} \gamma_{3s} FDI_{jt-1} \cdot Ds_{jt} \\ &+ \sum_{s=2}^{3} \gamma_{4s} PUB_{jt-1} \cdot Ds_{jt} + \sum_{s=2}^{3} \alpha_{js} \cdot Ds_{jt} + \sum_{s=2}^{3} \delta_{ts} \cdot Ds_{jt} + \epsilon_{jt} \end{aligned} \tag{4}$$

where the Marshallian externalities' variables are defined as above. Note that since the policy variables vary only by industry and year, we continue to interact them with the size indicators.

The last column of Table 5 presents the results from estimating Equation 5 while Table 6 presents the results from testing various hypotheses about the coefficient estimates. De-licensing reduces the spatial concentration of small plants ($\beta_1 < 0$) but raises concentration for medium-sized and large plants ($\beta_1 + \gamma_{12} > 0$ and $\beta_1 + \gamma_{13} > 0$) though these effects are insignificant. FDI liberalization reduces significantly the spatial concentration of small and

magnitudes.

medium-sized plants ($\beta_3 + \gamma_{32} < 0$ and $\beta_3 < 0$). The spatial concentration of large plants falls significantly in response to trade liberalization ($\beta_2 + \gamma_{23} > 0$) while that of small plants increases significantly ($\beta_2 < 0$). In industries with greater public sector presence large plants are more spatially concentrated ($\beta_4 + \gamma_{43} > 0$) while medium-sized plants are significantly less concentrated ($\beta_4 + \gamma_{42} < 0$).

8 Alternate Specifications

Tables 3 and 4 show that, among the traditional determinants of agglomeration, an industry's dependence on IO linkages and on transport increases spatial concentration in India. However, counterintuitively more labor pooling tends to lower spatial concentration. One concern is that a state might be too large a geographic unit to consider when measuring labor pooling. Plants may target smaller units, for example, districts with large proportions of educated or appropriately skilled labor force. However, when we estimate Equation 2 using district-level EG indices the result that greater labor pooling potential has a negative or zero effect on agglomeration is maintained. The counterintuitive result is also obtained for alternative measures of labor pooling including average employment shares of skilled workers and labor productivity in an industry.[35] The data show that the correlation between the EG index and labor pooling is large and negative initially, and then declines over time. Hence, there is a time dimension to the relationship between the demand for skill and geographic concentration. It is possible that as industries upgrade technology and skill-biased technical change occurs, labor pooling becomes an increasingly important determinant of agglomeration. Further, Chamarbagwalla and Sharma (2011) provide evidence of skill upgrading in Indian manufacturing during the 1980s and 1990s which would suggest that the effect of labor pooling on agglomeration may have changed from the 1980s to the 1990s.

Labor pooling is not the only factor whose relationship with agglomeration may change over time. As suggested by the discussion in Section 3.1, the negative effect of de-licensing on agglomeration could be due to the fact that de-licensing led to a break up of artificially created clusters in deemed backward areas. If this is the case, then we should see different trends in

[35]The results based on district-level EG indices and using alternative labor pooling measures are available upon request.

spatial concentration in the 1990s (when there were no restrictions on location of deregulated plants) compared to the 1980s (when plants in deregulated industries were deregulated only if they located in backward areas). Similarly as technologies and infrastructure improve over time, input-output linkages and transport dependence may affect agglomeration differentially.

Table B.6 shows that relative to the 1980s, de-licensing led to a rise in agglomeration during the 1990s.[36] This finding is consistent with the story that during the 1980s plants may have taken advantage of deregulation by setting up new establishments in backward areas where no clusters existed, leading to a decline in average agglomeration. During the 1990s plants begin to choose their locations optimally according to economic forces, for example seeking areas with good infrastructure, IO linkages, externalities and these areas tended to already have industrial clusters. We also find that labor pooling did not affect agglomeration in the 1990s and thus the negative effect of labor pooling on agglomeration is confined to the 1980s. This evidence is consistent with the story that as Indian industries upgraded technology, the benefits from labor pooling rose and caused these industries increasingly to agglomerate.

The relationship between the traditional and the policy determinants of geographic concentration may vary along an additional dimension: the trade orientation of the industry. For example, as trade reforms take effect and import-competing industries cut costs to compete with imports, they may become increasingly sensitive to agglomeration economies as a source of cost savings. Additionally, exporting industries may respond to de-licensing by relocating to geographic areas that reduce their costs and make them even more competitive in the world markets. In order to test these hypotheses, we calculate the trade orientation of each 3-digit industry and estimate Equation 2 for non-traded, import-competing and export-oriented industries separately.[37] To facilitate hypothesis testing we estimate also a fully interacted pooled model and present the results in Table B.7 defining $D2_{jt}$ to be the dummy for import-competing industries and $D3_{jt}$ the dummy for export-oriented industries.

[36] FDI liberalization is not included in this specification since that variable does not vary until 1990.

[37] We follow Pavcnik (2002) in defining industry trade orientation based on the average ratio of industry imports to output and of industry exports to output over the period 1980-99. An industry with an import-output ratio greater than 10% is considered import-competing. An industry with an export-output ratio greater than 10% is considered export-oriented. The remaining industries are categorized as non-traded. Our results are invariant to the use of thresholds other than 10%.

Table B.8 shows the results from testing various hypotheses about the coefficient estimates.

We find that de-licensing reduces significantly the geographic concentration of non-traded industries but the total effects on import-competing and export-oriented industries are insignificant. The only significant (and positive) effect of trade reforms on geographic concentration occurs, surprisingly, for non-traded industries. With respect to the traditional determinants of agglomeration, we find that most mechanisms are stronger for import-competing industries. In particular, import-competing industries agglomerate as a result of greater labor pooling potential while the non-traded industries disperse. This is consistent with the story that greater import competition forces import-competing industries to cut costs via greater use of localization economies.

9 Conclusion

In this paper we analyze the impact of industrial and trade policies on the agglomeration of manufacturing industries in India. Our results show the importance of policies - in particular de-licensing and FDI liberalization - in affecting the spatial distribution of Indian manufacturing. Our study is one of the few that uses a long time span of Indian plant-level data from 1980 to 1999 that allows us to consider these important changes in industrial policies.

Several of our results can inform policymakers, particularly in developing countries. Our results show that both de-licensing and FDI liberalization led to a decline in average agglomeration levels, while trade policy had no significant effect. That is, plants respond to some market-oriented policy reforms but not all. Further, our results point out the importance of an hitherto unexplored policy that can affect agglomeration – FDI liberalization – and emphasize the need to examine alternative mechanisms that affect plant location decisions.

Our results are also unique in that most studies for developed and developing countries alike have emphasized how market-based forces tend to increase agglomeration. Thus it is important to acknowledge the ambivalence of the theory and the prevalence of a multitude of mechanisms – including for example, the existing spatial distribution and congestion costs in existing clusters – that affect agglomeration, and hence should inform policy. Our result that policy reforms increase dispersion in Indian manufacturing provides an explanation for why

Bollard et al. (2010) find that only a quarter of the Indian aggregate productivity growth can be attributed to the policy reforms. Increased dispersion following the reforms reduced an important source of productivity gains - agglomeration economies.

Thirdly, our results emphasize the importance of accounting for heterogeneity across plants. We find considerable and consistent evidence that small, medium-sized and large plants respond differently to the policy determinants of agglomeration. Our results show the importance of looking beyond industry-level average effects. In the case of India we use plant size as the main source of plant-level heterogeneity because industrial policy in India pre-reform was conditional on plant size. However there are several other important dimensions along which plants are heterogenous and should be considered in future empirical studies as well as in policy decisions.

References

Amiti, M. Specialization patterns in Europe. *Weltwirschaftliches Archiv*, 134(4):573–593, 1999.

Amiti, M. and Javorcik, B. Trade Costs and Location of Foreign Firms in China. *Journal of Development Economics*, 85:129–149, 2008.

Amiti, M. and Cameron, L. Economic Geography and Wages. *The Review of Economics and Statistics*, 89(1):15–29, 2007.

Arrow, K. The Economic Implications of Learning By Doing. *The Review of Economics and Statistics*, 29:155–173, 1962.

Bai, C., Du, Y., Tao, Z., and Tong, S. Local Protectionism and Regional Specialization: Evidence from China's Industries. *Journal of International Economics*, 63:397–417, 2004.

Behrens, K., Gaigne, C., Ottaviano, G., , and Thisse, J. Countries, Regions and Trade: On the Welfare Impacts of Economic Integration. *European Economic Review*, 51:1277–1301, 2007.

Bollard, A., Klenow, P., and Sharma, G. India's Mysterious Manufacturing Miracle. *Mimeo*, September 2010.

Brülhart, M. The Spatial Effects of Trade Openness: A Survey. *Mimeo*, 2010.

Brülhart, M., Crozet, M., and Koenig, P. Enlargement and the EU Periphery: The Impact of Changing Market Potential. *The World Economy*, 27(6):853–875, 2004.

Brulhart, M. and Sbergami, F. Agglomeration and Growth: Cross-country Evidence. *Journal of Urban Economics*, 65:48–63, 2009.

Chamarbagwalla, R. and Sharma, G. Industrial Deregulation, Trade Liberalization, and Skill Upgrading in India. *Journal of Development Economics*, 96(2):314–336, 2011.

Ciccone, A. and Hall, R. Productivity and Density of Economic Activity. *American Economic Review*, 86(1):54–70, 1996.

Cingano, F. and Schivardi, F. Identifying the Sources of Local Productivity Growth. *Journal of the European Economic Association*, 2(4):720–742, 2004.

Combes, P.-P. and Overman, H. The Spatial Distribution of Economic Activities in the European Union. In Henderson, J. and Thisse, J., editors, *Handbook of Regional and Urban Economics*, volume 4. Elsevier: BV, 2004.

Crozet, M. and Koenig, P. EU Enlargement and the Internal Geography of Countries. *Journal of Comparative Economics*, 32:265–278, 2004.

Das, D. Manufacturing Productivity Under Varying Trade Regimes: India in the 1980s and 1990s. *ICRIER Working Paper 107*, 2003.

Deaton, A. and Dreze, J. Poverty and Inequality in India: A Re-examination. *Economic and Political Weekly*, pages 3729–3748, 7 September 2002.

Dixit, A. and Stiglitz, J. Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, 67(3):297–308, June 1977.

Dumais, G., Ellison, G., and Glaeser, E. Geographic Concentration as a Dynamic Process. *Review of Economics and Statistics*, 84(2):193–204, 2002.

Ellison, G. and Glaeser, E. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5):889–927, 1997.

Foster, L., Haltiwanger, J. C., and Krizan, C. J. Aggregate Productivity Growth. Lessons from Microeconomic Evidence. In *NBER Chapters*, chapter New Developments in Productivity Analysis, pages 303–372. NBER, 2001.

Fujita, M. and Hu, D. Regional Disparity in China 1985-1994: The Effects of Globalization and Economic Liberalization. *The Annals of Regional Science*, 35:3–37, 2001.

Ge, Y. Regional Inequality, Industry Agglomeration and Foreign Trade, The Case of China. *UNU-WIDER*, 2006/105, 2006.

Ge, Y. Globalization and Industry Agglomeration in China. *World Development*, 37(3): 550–559, 2009.

Glaeser, E., Kallal, H., Scheinkman, J., and Shleifer, A. Growth in Cities. *Journal of Political Economy*, 100:1126–1152, 1992.

Haaland, J., Kind, H.-J., and Knarvik, K.-H. M. What Determines the Economic Geography of Europe? *CEPR Working Paper No. 2072*, 1999.

Hanson, G. Localization Economies, Vertical Organization, and Trade. *American Economic Review*, 86(5):1266–1278, 1996.

Hanson, G. Increasing Returns, Trade, and Regional Structure of Wages. *Economic Journal*, 107(440):113–133, 1997.

Hanson, G. and Harrison, A. Trade and Wage Inequality in Mexico. *Industrial and Labor Relations Review*, 52(2):271–288, 1999.

Haskel, J., Pereira, S., and Slaughter, M. J. Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms? *The Review of Economics and Statistics*, 89(3):482–496, 2007.

Holmes, T. Localization Of Industry And Vertical Disintegration. *The Review of Economics and Statistics*, 81(2):314–325, 1999.

Jacobs, J. *The Economics of Cities.* Vintage New York, 1969.

Kim, S. Expansion of Markets and the Geographic Distribution of Economic Activities:The Trends in U.S. Regional Manufacturing Structure 1860-1987. *Quarterly Journal of Economics*, 110(4):881–908, 1995.

Krishna, P. and Mitra, D. Trade Liberalization, Market Discipline and Productivity Growth: New Evidence from India. *Journal of Development Economics*, 56:391–409, 1996.

Krugman, P. and Elizondo, R. L. Trade Policy and the Third World Metropolis. *Journal of Development Economics*, 49:137–150, 1996.

Krugman, P. *Geography and Trade.* Leuven University Press-MIT Press, 1991.

Krugman, P. and Venables, A. Globalization and the Inequality of Nations. *Quarterly Journal of Economics*, 60:857–880, 1995.

Lall, S. and Chakravorty, S. Industrial Location and Spatial Inequality: Theory and Evidence from India. *Review of Development Economics*, 9(1):47–68, 2005.

Lall, S. and Chakravorty, S. *Made in India.* Oxford University Press, 2007.

Lall, S., Shalizi, Z., and Deichmann, U. Agglomeration Economies and Productivity in Indian Industry. *Journal of Development Economics*, 73(2):643–673, 2004.

Lu, J. and Tao, Z. Trends and Determinants of China's Industrial Agglomeration. *Journal of Urban Economics*, 65:167–180, 2009.

Marathe, S. S. *Regulation and Development: India's Policy Experience of Controls over Industry.* SAGE Publications, second edition, 1989.

Marshall, A. *Principles of Economics.* MacMillan London, 1890.

Martincus, C. Spatial Effects of Trade Policy: Evidence from Brazil. *Journal of Regional Science*, 50(2):541–569, 2010.

Martincus, C. and Sanguinetti, P. Tariffs and Manufacturing Location in Argentina. *Regional Science and Urban Economics*, 39:155–167, 2009.

Mayer, K. and Mayer, T. The Empirics and Agglomeration and Trade. In Henderson, J. and Thisse, J., editors, *Handbook of Regional and Urban Economics*, volume 4. Elsevier: BV, 2004.

Mayer, T., Mejean, I., and Nefussi, B. The Location of Domestic and Foreign Production Affiliates by French Multinational Firms. *Journal of Urban Economics*, 68:115–128, 2010.

Monfort, P. and Nicolini, R. Regional Convergence and International Integration. *Journal of Urban Economics*, 48:286–306, 2000.

Ottaviano, G. and Thisse, J. Agglomeration and Foreign Direct Investment. In Rajan, R. and Reinert, K., editors, *Princeton Encyclopedia of the World Economy.* Princeton University Press, Princeton, N.J., 2008.

Overman, H. and Puga, D. Labour Pooling as a Source of Agglomeration: An Empirical Investigation. *CEPR Working Paper No. DP7174*, Feb 2009.

Pal, P. and Ghosh, J. Inequality in India: A Survey of Recent Trends. *DESA Working Paper*, (45), July 2007.

Paluzie, E. Trade Policy and Regional Inequalities. *Papers in Regional Science*, 80:67–85, 2001.

Pavcnik, N. Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants. *Review of Economic Studies*, (69), 2002.

Porter, M. *The Competitive Advantage of Nations.* New York Free Press, 1991.

Propris, L. D. and Driffield, N. The Importance of Clusters for Spillovers from Foreign Direct Investment and Technology Sourcing. *Cambridge Journal of Economics*, 30:277–291, 2006.

Romer, P. Increasing Returns and Longterm Growth. *Journal of Political Economy*, 94: 1002–1037, 1984.

Rosenthal, S. and Strange, W. The Determinants of Agglomeration. *Journal of Urban Economics*, 50(2):191–229, 2001.

Sharma, G. *Competing or Collaborating Siblings: An Investigation of the Relationship between Industrial and Trade Policies.* PhD thesis, University of Maryland, 2006a.

Sharma, G. Competing or Collaborating Siblings: Industrial and Trade Policies in India. *University of Missouri Working Paper 06-10*, 2006b.

Sivadasan, J. Barriers to Competition and Productivity: Evidence from India. *The B.E. Journal of Economic Analysis and Policy*, 9(1 (Advances)), 2009.

Topalova, P. Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty and Inequality from India. *MIT*, October 2004.

Topalova, P. B. and Khandelwal, A. Trade Liberalization and Firm Productivity: The Case of India. *Review of Economics and Statistics*, 93:995–1009, 2011.

Figures

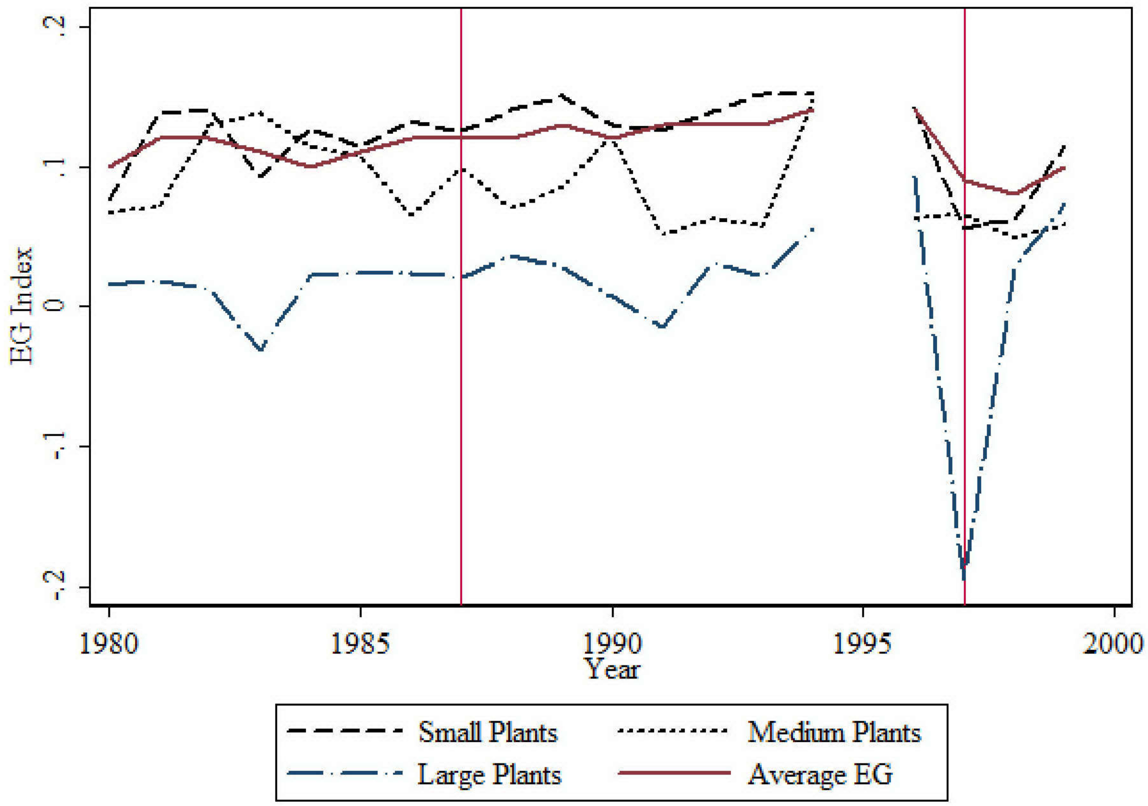

Figure 1: Weighted Average EG Index for different sized plants
Note: All averages are weighted by industry size (employment).

Tables

Table 1: Percentage of Employment, Output and Capital De-licensed in each year

year	Employment	Cumulative Employment	Cumulative Capital	Cumulative Output
1984	7.6	7.6	4.7	6.8
1985	18.3	15.0	11.7	13.9
1986	3.8	17.9	18.9	20.1
1987	26.5	23.5	24.3	25.0
1988		23.4	23.6	25.1
1989		22.4	21.2	22.8
1990		22.8	19.7	23.1
1991	60.0	90.0	90.2	84.2
1992		90.2	90.5	83.8
1993	2.6	90.9	90.5	84.8
1994		91.1	90.5	84.6
1996		99.5	99.6	99.7
1997		99.1	96.4	94.3
1998		73.0	66.9	67.6
1999		99.4	99.5	99.1

Note: To compute these figures we use de-licensing and plant-level data at the 4-digit level of NIC.

Table 2: Descriptive Statistics: EG Index for Indian Manufacturing

	Raw Concentration Index		EG Index				
	Mean	Standard Deviation	Mean	Standard Deviation	Min	Max	Total Number of industries
2-digit industries							
1980	0.086	0.127	0.076	0.142	-0.019	0.589	19
1990	0.08	0.134	0.083	0.146	0.003	0.618	19
1999	0.06	0.078	0.047	0.085	-0.073	0.368	20
3-digit industries							
1980	0.157	0.169	0.084	0.209	-0.870	1.045	162
1990	0.162	0.186	0.094	0.203	-1.013	0.942	174
1999	0.172	0.161	0.086	0.149	-0.149	1.097	175

Table 3: Estimation Results: Equation 2

Variable	Coefficient	Raw Concentration Index	EG Index	No IRS	Industry-Year FE
DEL_{jt-1}	β_1	-0.0249**	-0.0193**	-0.0184**	-0.0162*
		(0.0096)	(0.0076)	(0.0073)	(0.0098)
ERP_{jt-1}	β_2	0.0001	0.0036	0.0008	0.0016
		(0.0102)	(0.0083)	(0.0073)	(0.0084)
FDI_{jt-1}	β_3	-0.0308***	-0.0356***	-0.0380***	-0.0437***
		(0.0117)	(0.0123)	(0.0128)	(0.0160)
PUB_{jt-1}	β_4	0.0761***	0.0312	0.0041	0.0026
		(0.0278)	(0.0209)	(0.0181)	(0.0223)
$MATS_{jt-1}$	β_5	0.0217	0.0362	0.0540**	0.0578*
		(0.0251)	(0.0273)	(0.0250)	(0.0306)
$INVEN_{jt-1}$	β_6	0.0044	0.0022	0.0023	-0.0007
		(0.0029)	(0.0022)	(0.0021)	(0.0026)
$LPOOL_{jt-1}$	β_7	-0.1201***	-0.0977***	-0.1088***	-0.1480***
		(0.0419)	(0.0346)	(0.0351)	(0.0408)
IRS_{jt-1}	β_8	-0.0084*	-0.0087**		
		(0.0044)	(0.0039)		
Constant	α	0.0748	0.1082	0.0283	-0.0052
		(0.0787)	(0.0717)	(0.0600)	(0.0844)
No. of Observations	N	2264	2258	2258	2258
R-squared	R-sq	0.5845	0.5668	0.5639	0.7175

Note: Robust standard errors clustered at the 3-digit industry level in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. Year and 2-digit industry fixed effects are included. In Column 1 the dependent variable is the raw geographic concentration index. In Columns 2-4 the dependent variable is the EG index. All columns present the FGLS estimates where each observation is weighted by its employment size divided by its estimated variance from its industry-year mean ($w3$).

Table 4: Robustness Checks: Equation 2

Variable	Coeff-icient	Nominal Tariffs	Without trade	Without FDI	Exclude public sector reserved
DEL_{jt-1}	β_1	-0.0126	-0.0186***	-0.0170**	-0.0141
		(0.0099)	(0.0070)	(0.0074)	(0.0087)
TAR_{jt-1}	β_2	0.0100			
		(0.0216)			
PUB_{jt-1}	β_4	0.0195	0.0052	0.0047	
		(0.0195)	(0.0180)	(0.0179)	
FDI_{jt-1}	β_3	-0.0362***	-0.0373***		-0.0350***
		(0.0117)	(0.0126)		(0.0125)
$MATS_{jt-1}$	β_5	0.0493*	0.0520**	0.0555**	0.0542**
		(0.0277)	(0.0238)	(0.0251)	(0.0244)
$INVEN_{jt-1}$	β_6	0.0043*	0.0026	0.0026	0.0020
		(0.0026)	(0.0021)	(0.0021)	(0.0026)
$LPOOL_{jt-1}$	β_7	-0.1254***	-0.1073***	-0.1094***	-0.1117***
		(0.0365)	(0.0341)	(0.0352)	(0.0372)
ERP_{jt-1}	β_8			-0.0007	-0.0056
				(0.0075)	(0.0083)
Constant	α	0.0031	0.0322	0.0367	0.0482
		(0.0968)	(0.0526)	(0.0602)	(0.0610)
No. of Observations	N	1405	2287	2258	2152
R-squared	R-sq	0.5791	0.5617	0.5601	0.5676

Note: Robust standard errors clustered at the 3-digit industry level in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. The dependent variable is the EG index. Year and 2-digit industry fixed effects are included. All columns present the FGLS estimates where each observation is weighted by its employment size divided by its estimated variance from its industry-year mean ($w3$).

Table 5: Estimation Results: Equations 3 and 5

Variable	Coefficient	Small plants	Medium sized plants	Large plants	Pooled model
DEL_{jt-1}	β_1	-0.0241**	0.0196**	0.0358	-0.0446*
		(0.0120)	(0.0098)	(0.0221)	(0.0228)
ERP_{jt-1}	β_2	-0.0329**	-0.0018	0.0552**	-0.0595*
		(0.0132)	(0.0094)	(0.0214)	(0.0343)
PUB_{jt-1}	β_4	-0.0275	-0.0137	0.0489*	0.0276
		(0.0378)	(0.0182)	(0.0272)	(0.0753)
FDI_{jt-1}	β_3	-0.0530**	-0.0091	0.0241	-0.1262**
		(0.0247)	(0.0099)	(0.0221)	(0.0495)
$MATS^k_{jt-1}$	β_5	0.0689	-0.0073	-0.0426	0.0125
		(0.0438)	(0.0190)	(0.0453)	(0.0410)
$INVEN^k_{jt-1}$	β_6	0.0105	0.0053**	-0.0026	0.0008
		(0.0068)	(0.0026)	(0.0037)	(0.0031)
$LPOOL^k_{jt-1}$	β_7	-0.1596***	0.0102	-0.0469	-0.1549***
		(0.0512)	(0.0207)	(0.0407)	(0.0472)
$D2_{jt}$	β_{02}				-0.2588
					(0.1610)
$D3_{jt}$	β_{03}				-0.5526***
					(0.1684)
DEL_{jt-1}	γ_{12}				0.0601**
$*D2_{jt}$					(0.0243)
DEL_{jt-1}	γ_{13}				0.0785**
$*D3_{jt}$					(0.0336)
ERP_{jt-1}	γ_{22}				0.0565
$*D2_{jt}$					(0.0360)

continued on next page

Variable	Coefficient	Small plants	Medium sized plants	Large plants	Pooled model
ERP_{jt-1}	γ_{23}				0.1115***
$*D3_{jt}$					(0.0357)
FDI_{jt-1}	γ_{32}				0.0884*
$*D2_{jt}$					(0.0477)
FDI_{jt-1}	γ_{33}				0.1547***
$*D3_{jt}$					(0.0495)
PUB_{jt-1}	γ_{42}				-0.1118
$*D2_{jt}$					(0.0732)
PUB_{jt-1}	γ_{43}				0.0272
$*D3_{jt}$					(0.0792)
Constant	α	0.1780	0.0804	-0.3057***	0.2013
		(0.1280)	(0.0539)	(0.1153)	(0.1524)
No. of Observations		2277	1738	975	4990
R-squared		0.4810	0.7334	0.5760	0.5870

Note: Robust standard errors clustered at the 3-digit industry level in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. $D2_j = 1$ for medium-sized plants. $D3_j = 1$ for large plants. The dependent variable is the EG index. Year and 2-digit industry fixed effects are included. In Column 3 we include industry-size effects and year-size effects. All columns present the FGLS estimates where each observation is weighted by its employment size divided by its estimated variance from its industry-year mean ($w3$).

Table 6: Hypothesis Tests: Equation 5 with size varying regressors

Hypothesis	Estimate	S.E.	p-value
β_1	-0.04458	0.022823	0.052335
γ_{12}	0.060079	0.024273	0.01426
γ_{13}	0.078459	0.033591	0.020628
$\gamma_{12} - \gamma_{13}$	-0.01838	0.022744	0.420093
$\beta_1 + \gamma_{12}$	0.015495	0.013488	0.252193
$\beta_1 + \gamma_{13}$	0.033875	0.022307	0.130654
β_2	-0.05953	0.034257	0.084
γ_{22}	0.056532	0.035985	0.117975
γ_{23}	0.111452	0.035743	0.002125
$\gamma_{22} - \gamma_{23}$	-0.05492	0.024546	0.026505
$\beta_2 + \gamma_{22}$	-0.003	0.020126	0.881769
$\beta_2 + \gamma_{23}$	0.051922	0.021908	0.018862
β_4	0.027566	0.075301	0.71475
γ_{42}	-0.11182	0.073164	0.128201
γ_{43}	0.027205	0.07916	0.731497
$\gamma_{42} - \gamma_{43}$	-0.13903	0.056185	0.014286
$\beta_4 + \gamma_{42}$	-0.08426	0.046049	0.068971
$\beta_4 + \gamma_{43}$	0.054771	0.032085	0.089566
β_3	-0.12615	0.049491	0.011651
γ_{32}	0.088402	0.04766	0.065279
γ_{33}	0.154717	0.049507	0.002077
$\gamma_{32} - \gamma_{33}$	-0.06631	0.028927	0.023054
$\beta_3 + \gamma_{32}$	-0.03775	0.018933	0.047687
$\beta_3 + \gamma_{33}$	0.028562	0.02137	0.183075

A Appendix I: Indian industrial policy and geographic dispersion of industry

Marathe (1989) provides an in depth assessment of the objectives of the government and the way industrial policy was used to affect the dispersion of Indian industry. At the time of independence, manufacturing industry was highly localized. For example, out of 144 cotton mills, 100 were located in one district. The government was concerned about regional concentration of employment and income, as well as about social and economic costs associated with congestion. Further it was felt that 'new' enterprises may not choose their locations optimally since they might be guided by linguistic/regional preferences and political pressures. The government conceded that particularly in the early stages of development, access to infrastructure may increase agglomeration near urban areas (and that this was not necessarily a bad thing) and that even highly planned economies like the USSR had substantial variation in industrial distribution. But the Industrial Policy statements of 1945 and 1956 explicitly stated that a goal of industrial policy was to reduce concentration and/or develop industry in a list of "Industrially Backward States".

Marathe (1989) states that during the first 20 years (1950-70), the central government mainly used the distribution of public investment to try to affect the distribution of industry. More backward states (those with lower shares of employment in manufacturing and/or low per capita income) were assigned more investment. It was only in the 1970s that subsidy and concessional finance schemes started being used. Further, even within the backward states, some districts/areas were identified and even greater incentives were provided to locate in these areas. As a result there was a rise in the share of industry located in backward states but even within these states, there was an emergence of industrial areas near to already relatively developed industrial centers. Further, most of the concessional schemes significantly influenced the pattern of development of small enterprises, but not necessarily that of larger enterprises. From the 1970s onwards, the government also disproportionately located public sector enterprises in backward areas. In 1979, the share of industrially backward states in the output (employment) of centrally owned public enterprises was 61%(5[illegible]). But it is important to note that most of these figures are accounted for by states (which happen to be backward) with steel and coal mining operations.

The 1970s also saw the explicit use of the licensing regime as an important instrument for determining location decisions by giving positive weight to license applications that located new projects in backward areas. In the 1980s, the government streamlined and rationalized the subsidy and concessional finance schemes (offered by both the central and state governments). Further, they divided the country into three types of districts (based on presence of manufacturing industry), each with different amounts of subsidies and concessions. Other than stating that x% of subsidy/concession would be available if the project was located in area y, the exact choice was up to the entrepreneur. This discretion opened the door to the use of the licensing regime to locate industry in 'politically' popular areas. There were instances where the approval of the license was delayed until the entrepreneur was willing to move the location of the product to a backward district. In effect, since the licensing regime was being used explicitly to direct new investment to certain geographic areas it is important to analyze the long term effects, if any, of licensing.

B Appendix II: Supplementary Tables and Figures

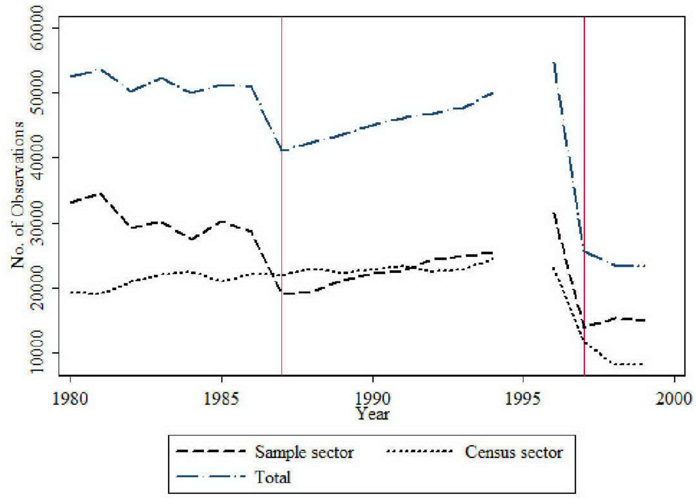

Figure B.1: Annual number of observations in ASI data

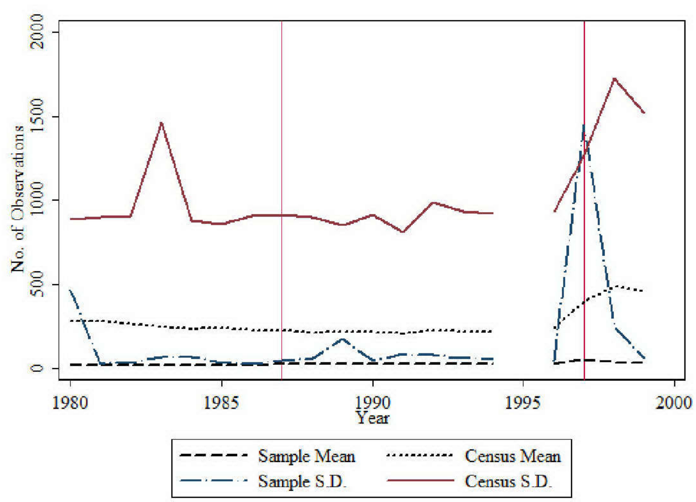

Figure B.2: Annual Average Employment and its Standard Deviation
Note: All averages are weighted by industry size (employment)

Table B.1: Annual Averages of Key Variables

Year	EG Index	$MATS_{jt}$	$INVEN_{jt}$	$LPOOL_{jt}$	IRS_{jt}	DEL_{jt}	PUB_{jt}	ERP_{jt}	FDI_{jt}
1980	0.10	0.57	-0.01	2085.28	2.64	0.00	0.33	90.36	0
1981	0.12	0.58	0.01	2371.42	2.54	0.00	0.31	90.45	0
1982	0.12	0.57	0.02	2566.29	3.31	0.00	0.30	91.28	0
1983	0.11	0.57	0.01	2636.69	2.97	0.00	0.31	91.63	0
1984	0.10	0.59	0.00	2954.33	3.17	0.19	0.33	93.96	0
1985	0.11	0.61	0.01	3248.50	2.78	0.23	0.32	93.24	0
1986	0.12	0.59	0.01	3382.43	2.71	0.27	0.32	101.02	0
1987	0.12	0.60	0.01	3517.64	2.38	0.31	0.32	101.71	0
1988	0.12	0.61	0.02	3915.60	2.71	0.31	0.32	101.26	0
1989	0.13	0.61	0.01	4237.95	2.77	0.28	0.25	103.58	0
1990	0.12	0.61	0.01	4616.20	3.11	0.29	0.25	104.53	0
1991	0.13	0.61	0.01	4450.14	2.87	0.96	0.24	60.79	0
1992	0.13	0.61	0.02	4646.35	3.25	0.96	0.22	61.40	0.22
1993	0.13	0.59	0.01	4872.91	3.35	0.97	0.18	61.30	0.22
1994	0.14	0.60	0.01	5217.37	3.81	0.97	0.15	61.48	0.22
1996	0.14	0.73	0.01	5979.65	4.32	0.99	0.15	32.21	0.23
1997	0.09	0.60	0.01	5688.28	4.17	0.99	0.17	35.31	0.19
1998	0.08	0.58	0.00	12546.07	4.13	0.73	0.14	34.09	0.20
1999	0.10	0.59	0.01	16753.86	3.97	0.99	0.13	32.03	0.24

Note: All averages are weighted by industry size (employment). The definition of the variables are as follows. $MATS_{jt}$ = Materials per Rs. Sales; $INVEN_{jt}$ = Inventories per Rs. Sales; $LPOOL_{jt}$ = Rs. sales per employee; IRS_{jt} = Real capital per plant ('00000 Rupees); DEL_{jt} = Proportion of Output De-licensed; PUB_{jt} = Proportion of Output under Public Sector; ERP_{jt} = Effective Rate of protection; FDI_{jt} = Proportion of Output FDI liberalized

Table B.2: Most Geographically Concentrated Industries in 1980

NIC	Description	EG	Rank in 1980	Rank in 1990	Rank in 1994	Rank in 1999
318	Manufacture of coke oven products	1.044525	1	1	1	3
229	Manufacture of pan masala, catechu and chewing lime	0.772687	2	20	26	62
325	Manufacture of mica products	0.760[illegible]5	3	2	11	167
225	Preparation of raw tobacco	0.724786	4	6	4	8
307	Manufacture of matches	0.640719	5	3	3	155
254	Spinning, weaving and finishing of jute and mesta textiles	0.608692	6	5	2	166
215	Processing of edible nuts	0.560939	7	10	9	67
214	Coffee curing, roasting, grinding and blending	0.47[illegible]5[illegible]8	8	12	14	7
385	Manufacture of sports and athletic goods	0.468759	9	39	34	43
213	Processing and blending of tea	0.434877	10	11	13	26

Table B.3: Least Geographically Concentrated Industries in 1980

NIC	Description	EG	Rank in 1980	Rank in 1990	Rank in 1994	Rank in 1999
227	Manufacture of cigars, cigarette, cheroots and cigarette tobacco	-0.04229	153	134	152	56
279	Manufacture of products of wood, bamboo, cane, reed, grass n.e.c.	-0.04525	154	66	139	161
275	Manufacture of cork and cork products	-0.10486	155	158	119	124
378	Manufacture of bullock-carts, push-carts and hand-carts	-0.14342	156	27	167	84
342	Manufacture of furniture and fixtures primarily of metal	-0.15044	157	147	163	162
384	Minting of currency coins	-0.34269	158	172	171	19
234	Weaving and finishing of cotton textiles on powerloom	-0.35382	159	77	71	133
257	Bleaching, dyeing and printing of just and mesta textiles	-0.35834	160	170	170	4
367	Manufacture of computers and computer based systems	-0.76451	161	156	70	93
386	Manufacture of musical instruments	-0.86979	162	168	168	2

Table B.4: Correlation Matrix for Key Variables

	EG Index	$MATS_{jt}$	$INVEN_{jt}$	$LPOOL_{jt}$	IRS_{jt}	DEL_{jt}	PUB_{jt}	ERP_{jt}	FDI_{jt}
EG Index	1								
$MATS_{jt}$	0.040	1							
$INVEN_{jt}$	0.040	0.009	1						
$LPOOL_{jt}$	-0.077	0.025	-0.022	1					
IRS_{jt}	-0.059	0.047	-0.017	0.634	1				
DEL_{jt}	-0.020	-0.007	-0.026	0.124	-0.020	1			
PUB_{jt}	-0.025	0.035	-0.046	0.293	0.184	0.347	1		
ERP_{jt}	-0.086	-0.003	0.022	-0.090	-0.011	-0.370	-0.347	1	
FDI_{jt}	-0.0563	0.0006	-0.0157	0.0331	-0.031	0.2185	-0.0368	-0.0892	1

Note: The definition of the variables are as follows. $MATS_{jt}$ = Materials per Rs. Sales; $INVEN_{jt}$ = Inventories per Rs. Sales; $LPOOL_{jt}$ = Rs. sales per employee; IRS_{jt} = Real capital per plant; DEL_{jt} = Proportion of Output De-licensed; PUB_{jt} = Proportion of Output under Public Sector; ERP_{jt} = Effective Rate of protection; FDI_{jt} = Proportion of Output FDI liberalized.

Table B.5: Estimation Results for Equation 2 with various weights

Variable	Wt = L_{jt}	Weight = W1	Weight = W2	Lag=2	Lag=3
DEL_{jt-k}	-0.0163	-0.0182**	-0.0166	-0.0274***	-0.0248**
	(0.0149)	(0.0072)	(0.0148)	(0.0103)	(0.0101)
ERP_{jt-k}	-0.0174	-0.0008	-0.0195	0.0079	0.0081
	(0.0166)	(0.0069)	(0.0180)	(0.0096)	(0.0100)
FDI_{jt-k}	-0.0385*	-0.0357**	-0.0424**	-0.0199	-0.0250
	(0.0222)	(0.0142)	(0.0213)	(0.0137)	(0.0166)
$MATS_{jt-k}$	0.0927**	0.0613***	0.0759	0.0669*	0.0631*
	(0.0432)	(0.0227)	(0.0500)	(0.0342)	(0.0325)
$INVEN_{jt-k}$	0.0096**	0.0029	0.0087*	-0.0015	-0.0030
	(0.0045)	(0.0021)	(0.0046)	(0.0023)	(0.0026)
$LPOOL_{jt-k}$	-0.1913***	-0.1066***	-0.1981***	-0.0622***	-0.0601***
	(0.0546)	(0.0340)	(0.0570)	(0.0169)	(0.0172)
PUB_{jt-k}	-0.0118	0.0103	-0.0168	0.0017	0.0031
	(0.0394)	(0.0181)	(0.0394)	(0.0217)	(0.0214)
Constant	0.0476	0.0343	0.0443	0.5972***	0.5647***
	(0.1028)	(0.0556)	(0.1109)	(0.1549)	(0.1560)
No. of Observations	2258	2258	2258	2172	2052
R-squared	0.5744	0.5276	0.6177	0.5479	0.5454
Breusch-Pagan test Statistic	8.61				
White's test Statistic	1195				

Note: Robust standard errors clustered at the 3-digit industry level in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. The dependent variable is the EG index. Year and 2-digit industry fixed effects are included. FGLS estimates where each observation is weighted by: its employment size are presented in Column 1, its employment size divided by its estimated variance from the industry-mean in Column 2; its employment size divided by its estimated variance from the annual mean in Column 3. Columns 4 and 5 present the FGLS estimates where each observation is weighted by its employment size divided by its estimated variance from the industry-year mean ($w3$).

Table B.6: Estimation Results: Differences Post 1991

Variable	Coefficient	1980-90	1991-99	Pooled model
DEL_{jt-1}	β_1	-0.0170**	0.0141*	-0.0185
		(0.0077)	(0.0084)	(0.0172)
ERP_{jt-1}	β_2	0.0045	0.0009	-0.0086
		(0.0072)	(0.0074)	(0.0125)
PUB_{jt-1}	β_4	-0.0042	0.0233	-0.0077
		(0.0176)	(0.0200)	(0.0391)
$MATS_{jt-1}$	β_5	0.0575**	0.0580**	0.0727
		(0.0249)	(0.0260)	(0.0531)
$INVEN_{jt-1}$	β_6	-0.0002	0.0039	0.0066
		(0.0019)	(0.0026)	(0.0060)
$LPOOL_{jt-1}$	β_7	-0.1268***	-0.1274***	-0.1766**
		(0.0422)	(0.0368)	(0.0693)
$D2_t$	β_{02}			-0.0409
				(0.0660)
DEL_{jt-1}	γ_{12}			0.0276
$*D2_{jt}$				(0.0217)
$MATS_{jt-1}$	γ_{52}			0.0311
$*D2_{jt}$				(0.0461)
$INVEN_{jt-1}$	γ_{62}			0.0027
$*D2_{jt}$				(0.0074)
$LPOOL_{jt-1}$	γ_{72}			-0.0358
$*D2_{jt}$				(0.0378)
Constant		-0.0102	0.0066	0.0144
		(0.0556)	(0.0618)	(0.1168)

continued on next page

Variable	Coefficient	1980-90	1991-99	Pooled model
Observations		1247	1011	2258
R-squared		0.7346	0.5502	0.5752

Note: Robust standard errors clustered at the 3-digit industry level in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. $D2_t = 1$ if year>1990. The dependent variable is the EG index. 2-digit industry fixed effects are included. All columns present the FGLS estimates where each observation is weighted by its employment size divided by its estimated variance from its industry-year mean ($w3$).

Table B.7: Estimation Results: Trade Orientation

Variable	Coefficient	Import Competing	Export Oriented	Non-traded	Pooled model
DEL_{jt-1}	β_1	0.0020	-0.0411	-0.0266*	-0.0309**
		(0.0057)	(0.0360)	(0.0151)	(0.0150)
ERP_{jt-1}	β_2	-0.0202	-0.0353	-0.0352***	-0.0356***
		(0.0192)	(0.0556)	(0.0109)	(0.0115)
FDI_{jt-1}	β_3	-0.0001	-0.5448	-0.0158	-0.0214
		(0.0102)	(0.6305)	(0.0196)	(0.0185)
PUB_{jt-1}	β_4	-0.0249	-0.3348	-0.0295	-0.0304
		(0.0210)	(0.2093)	(0.0180)	(0.0187)
$MATS_{jt-1}$	β_5	0.0696*	-0.0712	0.0728*	0.0830**
		(0.0346)	(0.0841)	(0.0398)	(0.0414)
$INVEN_{jt-1}$	β_6	0.0020	0.0042	0.0099***	0.0106***
		(0.0025)	(0.0065)	(0.0032)	(0.0034)
$LPOOL_{jt-1}$	β_7	0.0669**	-0.0174	-0.1907***	-0.1917***
		(0.0327)	(0.1284)	(0.0222)	(0.0241)
$D2_j$	β_{02}				0.1819*
					(0.1082)
$D3_j$	β_{03}				0.0923
					(0.2116)
DEL_{jt-1}	γ_{12}				0.0300*
$*D2_{jt}$					(0.0165)
DEL_{jt-1}	γ_{13}				-0.0074
$*D3_{jt}$					(0.0340)
ERP_{jt-1}	γ_{22}				-0.0065
$*D2_{jt}$					(0.0224)

continued on next page

Variable	Coefficient	Import Competing	Export Oriented	Non-traded	Pooled model
ERP_{jt-1}	γ_{23}				0.0050
$*D3_{jt}$					(0.0538)
FDI_{jt-1}	γ_{32}				-0.0327
$*D2_{jt}$					(0.0420)
FDI_{jt-1}	γ_{33}				-0.5618
$*D3_{jt}$					(0.6188)
PUB_{jt-1}	γ_{42}				0.0476*
$*D2_{jt}$					(0.0269)
PUB_{jt-1}	γ_{43}				-0.3166
$*D3_{jt}$					(0.2147)
$MATS_{jt-1}$	γ_{52}				0.0733
$*D2_{jt}$					(0.0550)
$MATS_{jt-1}$	γ_{53}				-0.1579
$*D3_{jt}$					(0.0966)
$INVEN_{jt-1}$	γ_{62}				-0.0133***
$*D2_{jt}$					(0.0051)
$INVEN_{jt-1}$	γ_{63}				-0.0031
$*D3_{jt}$					(0.0069)
$LPOOL_{jt-1}$	γ_{72}				0.2404***
$*D2_{jt}$					(0.0419)
$LPOOL_{jt-1}$	γ_{73}				0.1694
$*D3_{jt}$					(0.1400)
Constant	α	0.2374**	0.2490	0.1310*	0.1427**
		(0.0940)	(0.2274)	(0.0661)	(0.0695)
No. of Observations		430	533	1295	2258
R-squared		0.22[illegible]	0.2005	0.4278	0.3985

continued on next page

Variable	Coefficient	Import Competing	Export Oriented	Non-traded	Pooled model

Note: Robust standard errors clustered at the 3-digit industry level in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. $D2_j = 1$ if industry is import-competing. $D3_j = 1$ if industry is export-oriented. The dependent variable is the EG index. Year fixed effects are included. All columns present the FGLS estimates where each observation is weighted by its employment size divided by its estimated variance from its industry-year mean ($w3$).

Table B.8: Hypothesis Tests: Trade Orientation

Hypothesis	Estimate	S.E.	p-value	Hypothesis	Estimate	S.E.	p-value
β_1	-0.031	0.015	0.041	β_5	0.083	0.041	0.046
γ_{12}	0.030	0.016	0.071	γ_{52}	0.073	0.055	0.185
γ_{13}	-0.007	0.034	0.828	γ_{53}	-0.158	0.097	0.104
$\gamma_{12} - \gamma_{13}$	0.037	0.031	0.234	$\gamma_{52} - \gamma_{53}$	0.231	0.095	0.015
$\beta_1 + \gamma_{12}$	-0.001	0.007	0.897	$\beta_5 + \gamma_{52}$	0.156	0.036	0.000
$\beta_1 + \gamma_{13}$	-0.038	0.031	0.212	$\beta_5 + \gamma_{53}$	-0.075	0.087	0.392
β_2	-0.036	0.011	0.002	β_6	0.011	0.003	0.002
γ_{22}	-0.006	0.022	0.774	γ_{62}	-0.013	0.005	0.010
γ_{23}	0.005	0.054	0.926	γ_{63}	-0.003	0.007	0.652
$\gamma_{22} - \gamma_{23}$	-0.011	0.056	0.838	$\gamma_{62} - \gamma_{63}$	-0.010	0.007	0.156
$\beta_2 + \gamma_{22}$	-0.042	0.019	0.031	$\beta_6 + \gamma_{62}$	-0.003	0.004	0.489
$\beta_2 + \gamma_{23}$	-0.031	0.053	0.561	$\beta_6 + \gamma_{63}$	0.008	0.006	0.214
β_4	-0.030	0.019	0.107	β_7	-0.192	0.024	0.000
γ_{42}	0.048	0.027	0.078	γ_{72}	0.240	0.042	0.000
γ_{43}	-0.317	0.215	0.142	γ_{73}	0.169	0.140	0.228
$\gamma_{42} - \gamma_{43}$	0.364	0.215	0.092	$\gamma_{72} - \gamma_{73}$	0.071	0.142	0.618
$\beta_4 + \gamma_{42}$	0.017	0.019	0.372	$\beta_7 + \gamma_{72}$	0.049	0.034	0.157
$\beta_4 + \gamma_{43}$	-0.347	0.214	0.106	$\beta_7 + \gamma_{73}$	-0.022	0.138	0.872
β_3	-0.021	0.018	0.248				
γ_{32}	-0.033	0.042	0.437				
γ_{33}	-0.562	0.619	0.365				
$\gamma_{32} - \gamma_{33}$	0.529	0.620	0.394				
$\beta_3 + \gamma_{32}$	-0.054	0.038	0.153				
$\beta_3 + \gamma_{33}$	-0.583	0.619	0.347				