Inequality and Demand-Driven Innovation: Evidence from International Patent Applications

[HIGHLY PRELIMINARY; PLEASE ONLY CITE WITH AUTHOR CONSENT]

You can download the most current version here.

Sabrina Dorn*

Christian Kiedaisch†

March 20, 2017

Abstract

This paper studies how the distribution of income across consumers affects innovation by affecting the demand for new goods. Within a model with non-homothetic preferences, we show that inequality is more likely to be harmful for innovation when innovations become more incremental, but that it is more likely to be beneficial when the size of the population is increased. The model is extended to a multi-country setting in which it is shown that inequality affects the number of patent flows (applications of patents that are already granted elsewhere) towards a country in the same way as it affects innovation. In an empirical analysis based on a large panel data set from PATSTAT, we find that inequality is more likely to increase and less likely to decrease international patent flows towards a country the larger the size of the population and the lower GDP of the country is. These results are in line with the model predictions and robust to the inclusion of many control variables.

*me@sabrinadorn.com

†University of Zürich, christian.kiedaisch@econ.uzh.ch, I gratefully acknowledge financial support by the Swiss National Science Foundation

1 Introduction

This paper analyzes, both from a theoretical and empirical perspective, how the distribution of income across consumers affects innovation by affecting the demand for new goods. Innovation is the main driving force of long run growth and most governments are heavily involved in redistributing income in order to reduce the level of inequality. With income inequality on the rise in many developed countries (see for example Piketty, 2014), it is therefore important to understand how inequality and innovation are related.

The motivation for this paper is that rich and poor households differ with respect to their consumption pattern and that rich households consume a larger variety of goods than poorer ones. This is for example shown in empirical studies conducted by Jackson (1984) and Falkinger and Zweimüller (1996). Using data from the US consumer expenditure survey, Kiedaisch (2016) shows that there is also a strong positive association between a more narrowly defined variety of "innovative" goods purchased by a household and household expenditures. When a new innovative good is more likely to be purchased by a rich household, the demand for it and the incentives to invent it therefore depend on the distribution of income.

In order to better understand the effect of inequality on innovation, we analyze a product variety model based on Föllmi and Zweimüller (2016) in which households either consume or do not consume a particular good, and in which rich households consume a larger variety of goods than poorer households. There are fixed costs of innovating and innovations lead to cost reductions that transform competitively supplied non-innovative into monopolistically supplied innovative goods. Households differ with respect to income and there is free entry into R&D. Within this setting, the incentives to innovate depend on the number of households that purchase an innovative good and also on their willingness' to pay for the good. In equilibrium, poor households spend all their income on innovative mass consumption goods while rich households also purchase some luxurious non-innovative (service) goods (in Appendix A2, an extension is analyzed in which all households first purchase some non-innovative basic need goods like food before they start purchasing innovative goods and non-innovative luxuries. As long as most households are rich enough to purchase all basic need goods, the results remain the same).

The prices of innovative goods are inversely related to the number of households that purchase them in equilibrium. This implies that progressive transfers among poor households who spend all their income on innovative goods shift demand towards sectors with lower markups and thereby reduce the incentives to innovate and the number of innovations that are carried out. An reduction in inequality can therefore stiffle innovation through such price effects. When, however, income is redistributed from a rich household who spends additional income on non-innovative luxuries to a poorer "mass consumer", this increases the demand for innovative goods and the incentives to innovate.

Whether or not a progressive transfer between two randomly drawn households increases or decreases innovation depends on the endogenous fraction of households that spend all their income on innovative goods. This fraction increases when the size of the population increases (for a given total income), when innovations lead to larger cost reductions, and when total income decreases. The reasons for this are the following: due to the presence of substitutable non-innovative goods, innovators are restricted in their price setting power and can maximally charge a limit price. When the size of the population increases, an innovator selling at the limit price can break even by selling one unit of the innovative good to a smaller fraction of the population. When cost reductions are larger, the limit price increases and innovators can also break even by selling to a smaller fraction of the population. In both cases, the fraction of households that only consume innovative goods in equilibrium increases. When total income increases, overall more innovations are carried out and the marginal innovation is of lower value. This implies a lower limit price, forcing innovators to sell to a larger fraction of the population in order to break even. Taken together, these effects imply that a progressive (inequality-reducing) transfer between two randomly drawn households is less likely to encourage or more likely to discourage innovation the larger the size of the population is, the larger the size of the cost reductions of innovations are, and the lower total income is.

The model is extended to a multi-country setting with two types of fixed costs: Fixed costs of invention (that are borne only once), and country-specific fixed costs of (subsequent) patenting or technology adoption. In this setting, inventors only pay the fixed costs of patenting or adoption when there is sufficient demand for an innovative good in a particular country (it is assumed that there is no parallel trade). In equilibrium, some countries endogenously emerge as "frontier" countries in which R&D is undertaken and others are "follower" countries in which only a fraction of the global stock of inventions get patented. When an innovation undertaken and patented in one country also gets subsequent patent protection in another country, a patent flow occurs between the two countries. The model predicts that inequality in a follower country is more likely to be good and less likely to be bad for the number of international patent flows towards this country the larger the size of the population, the larger the cost saving of the innovation, and the lower total income in this country is.

The qualitative (interaction) effects that inequality has on international patent flows towards a country are therefore the same as those that inequality has on innovation in the closed economy model. By empirically analyzing how inequality affects international patent flows we therefore hope to be able to infer how inequality affects innovation, at least in a qualitative way.

From an empirical perspective, studying patent flows has the following advantages: while the incentives to innovate might be driven by the demand coming from several countries and not only the country of invention, a patent flow towards a single country usually only occurs if the profitability in this country is sufficiently large. There is therefore a clearer definition of the relevant market. Moreover, we can study the variation in patent flows originating from the same set of technologies and do not need to compare heterogeneous innovations across countries. Furthermore, we believe that it is unlikely that a patent flow towards a country affects the level of inequality there (reverse causality) as technologies might diffuse towards countries independently of whether they are (subsequently) patented there (see Appendix A3).

Based on raw-data from the PATSTAT data-base on patents of invention in the manufacturing sector within 1980-2013, we study the behavior of international patent-flows from a country of origin (frontier-country) to a different country (adopting-country) with respect to the implications of the claimed theoretical channel.

Using a rich set of control variables, and considering five different measures of inequality (Gini coefficients pre- and post-taxes as well as top income shares) yielding to estimation samples of approximately 1,8 - 2,9 million observations, we find a strong statistical pattern supporting the theoretical predictions. This is confirmed by four different estimators of the conditional expectation of patent-flows parametrized by an exponential mean model.

For a quantitative assessment of our estimation results, using a sample of 67 countries from 1980-2013, we may predict the incidence of a positive marginal effect of an increase in inequality for all empirically observed combinations of population and GDP over these countries as destinations and all points in time. We estimate the probabilities that an increase in inequality increases patent flows to be roughly 0.57 when inequality is measured in terms of Gini post-tax, 0.53 when it is measured in terms of Gini pre-tax, and 0.26, 0.23, and 0.11 when it is measured in terms of the top-10%, top-5%, and top-1% income share.

In an extension of the model, we find that strengthening patent protection makes it more likely that inequality increases (and less likely that it decreases) innovation. Measuring the strength of patent protection by the Ginarte-Park index, we indeed find this predicted positive interaction between patent protection and the effect of inequality on patent flows in our empirical analysis.

2 Related Literature (to be added)

The theoretical model builds on the analysis of Föllmi and Zweimüller (2016) who analyze the effect of inequality on growth in an endogenous growth model. The present model is static (i.e. there are only two periods), but, unlike Föllmi and Zweimüller (2016) considers the more general case where the size of the population can vary and where the value of the marginal innovation can depend on the number of innovations that have already been carried out. While Föllmi and Zweimüller merely show that inequality can be either beneficial or harmful for growth (depending on whether price or market size effects dominate), the present model shows that inequality is more likely good or less likely bad for innovation the larger the size of the population and the lower total income is. Moreover, we extend the model to an international setup in order to derive empirically testable implications, while Föllmi and Zweimüller (2016) only consider the case of a closed economy and do not undertake any empirical analysis.

3 Model

In this section we develop the theoretical model.

3.1 Preferences

All households have the same utility function given by

$$U = f(C) + D \tag{1}$$

where $C = \int_{j=0}^{N} c_j dj$ denotes the variety of innovative goods and $D = \int_{n=0}^{\infty} c_n dn$ the variety of non-innovative goods consumed. Goods are indivisible and households are assumed to be satiated with a particular good after consuming one unit of it. Therefore, households either consume one or zero units of any good (i.e. $c_j \in \{0; 1\}$ and $c_n \in \{0; 1\}$). While there is an infinite variety of producible non-innovative goods[1], only the endogenous measure N of innovative goods can be produced. It is assumed that $f(C)$ is a continuous function satisfying $\frac{\partial f(C)}{\partial C} > 0$, $\frac{\partial^2 f(C)}{\partial C^2} < 0$, $\left.\frac{\partial f(C)}{\partial C}\right|_{C=0} = 1$ and $\lim\limits_{C\to\infty} \frac{\partial f(C)}{\partial C} > \frac{c}{\Omega}$, where $0 < \frac{c}{\Omega} < 1$ (the latter is an assumption on cost parameters which is explained further below). These assumptions imply that the utility derived from consuming an additional innovative good lies (weakly) below that of consuming an additional non-innovative good (which is constant) and that the former falls in the measure C of innovative goods that

[1] The analysis would be the same if there was instead only one divisible non-innovative good the quantity of which entered linearly into household's utility.

a household consumes[2]. The more innovative goods a household consumes the less it therefore values each single one of them relative to a non-innovative good. The analysis is, however, also carried out for the case in which the marginal utility from consuming innovative goods is constant, i.e. in which $\frac{\partial^2 f(C)}{\partial C^2} = 0$ and $\frac{\partial f(C)}{\partial C} = 1$ holds for all values of C).

3.2 Endowments, technology, and market structure

Labor is the only production factor. The total labor endowment in terms of efficiency units is given by Y and the size of the population by L. Households are assumed to differ with respect to their labor productivities and a household of type θ is endowed with $y(\theta) = \theta \frac{Y}{L}$ efficiency units of labor. The type θ therefore indicates the labor endowment of the household divided by the average labor endowment in the economy (both measured in efficiency units). The variable θ is distributed with density $g(\theta)$ and cdf $G(\theta)$ in the interval between $\underline{\theta} < 1$ and $\bar{\theta} > 1$. As $\theta = y(\theta)\frac{L}{Y}$ and $Y = L\int_{\underline{\theta}}^{\bar{\theta}} y(\theta)g(\theta)d\theta$, the condition $\int_{\underline{\theta}}^{\bar{\theta}} \theta g(\theta)d\theta = 1$ must hold.

For a given distribution $G(\theta)$ an increase in Y implies that all labor endowments $y(\theta)$ increase proportionally (i.e. an increasing scale transformation), while an increase in the size of the population L implies that all labor endowments are reduced proportionally (i.e. a decreasing scale transformation).

In order to produce one unit of a non-innovative good, Ω units of labor are required, while only $c < \Omega$ units are required in order to produce one unit of an innovative good. While the technology to produce non-innovative goods is in the public domain, innovative (low-cost) goods first need to be invented: The fixed costs F in terms of labor need to be paid in order to invent one innovative good j.

This set-up might be interpreted in the following way: Innovations transform non-innovative goods (e.g. traditional manufacturing goods or personal services) into innovative goods (e.g. modern manufacturing goods) that can be supplied at lower marginal cost (due to the use of automation). The assumption that $f(C)$ is a concave function is supposed to capture effects that might also arise in a more complicated (and less tractable) setup in which innovations are heterogenous with respect to their cost saving potential and in which the cost reduction of the marginal innovation falls in the measure N of innovations that are carried out.

Innovators obtain patents on their inventions that allow them to exclude others from using their technologies. Labor markets are assumed to be competitive and the wage rate for one efficiency unit of labor is normalized to one. There is free entry into R&D.

[2]Innovative goods are therefore only consumed if their price is sufficiently low (see below).

3.3 Consumption choices

As the blueprints for non-innovative goods are in the public domain, they are sold at the marginal cost $p_n = \Omega$. Denoting the (per unit) price of innovative good j by p_j, the budget constraint of a household with income y is given by

$$y = \int_{j=0}^{N} c_j(y) p_j dj + \int_{n=0}^{\infty} c_n(y) \Omega dn \tag{2}$$

with the left hand side denoting income[3] and the right hand side the expenditures on innovative (first term) and non-innovative (second term) goods. Maximizing utility (equation (1)) subject to this budget constraint leads to the following optimal consumption rules of a household with income y, where the Lagrange multiplier $\lambda(y)$ denotes the marginal utility of income and $z_j(y)$ ($z_n(y)$) the willingness to pay for an innovative (non-innovative) good of the corresponding household:

$$c_j(y) = \begin{cases} 1 & \text{if} \quad p_j \leq \left(\frac{1}{\lambda(y)}\right) \frac{\partial f(C(y))}{\partial C(y)} \equiv z_j(y) \\ 0 & \text{if} \quad p_j > z_j(y) \end{cases}$$

$$c_n(y) = \begin{cases} 1 & \text{if} \quad p_n = \Omega < \frac{1}{\lambda(y)} \equiv z_n(y) \\ 1 \text{ or } 0 & \text{if} \quad \Omega = z_n(y) \\ 0 & \text{if} \quad \Omega > z_n(y). \end{cases}$$

Households therefore consume a good when their willingness to pay exceeds (or is equal to) its price and do not consume it otherwise. While $z_j(y)$ and $z_n(y)$ coincide when a household does not consume any innovative good (as $z_j(y) = z_n(y) \frac{\partial f(C(y))}{\partial C(y)} = z_n(y)$ when $C(y) = 0$ due to the assumption that $\left.\frac{\partial f(C)}{\partial C}\right|_{C=0} = 1$), the willingness to pay for innovative goods ($z_j(y)$) falls short of that for non-innovative goods ($z_n(y)$) once a household consumes a positive variety $C(y)$ of innovative goods, and the more so, the larger this variety is (this is because $\frac{\partial f(C)}{\partial C}$ falls in C). Due to the assumption that $\lim_{C\to\infty} \frac{\partial f(C)}{\partial C} > \frac{c}{\Omega}$, households always prefer innovative to non-innovative goods when both are sold at marginal cost, i.e. when $p_n = \Omega$ and when $p_j = c$. In equilibrium, $z_j(y)$ and $z_n(y)$ weakly increase in income y and rich households consume some goods that poorer

[3] Due to free entry into R&D, net profits of innovators are equal to zero, implying that labor is the only source of income.

households do not consume.

The inventor of innovative good j can only earn positive profits and break even if the price p_j lies above the marginal production costs c. Given that all innovative goods are sold at prices that are low enough for consumers to prefer them to non-innovative goods (this must be the case in equilibrium), households first consume the cheapest innovative goods and then spend their incremental income on more and more expensive innovative goods (in equilibrium, these goods are sold at different prices when households differ with respect to the variety of innovative goods that they consume). Only households with $y > \hat{y}$ who are rich enough to purchase one unit of each of the N innovative goods also purchase some non-innovative goods (that are sold at price Ω). The varieties $C(y)$ of innovative goods and $D(y)$ of non-innovative goods consumed by a household with income y are therefore given by:

$$C(y) = \begin{cases} \int_{j=0}^{N} c_j(y) dj & \text{if} \quad y \leq \hat{y} \\ N & \text{if} \quad y > \hat{y} \end{cases}$$

$$D(y) = \begin{cases} 0 & \text{if} \quad y \leq \hat{y} \\ \frac{y - \int_{j=0}^{N} p_j dj}{\Omega} & \text{if} \quad y > \hat{y} \end{cases}$$

When income lies below the threshold $\hat{y}$, $C(y)$ rises in household income, so that, in line with the empirical evidence, richer households purchase a larger variety of innovative goods than poorer ones.

In Appendix A2 an extension is discussed in which there are some basic need goods (like food) which households purchase before they start spending money on innovative and other non-innovative (service) goods. The qualitative results of the analysis remain similar in such a more general setup.

3.4 Equilibrium price structure

An innovator selling innovative good j sets the price $p(j)$ in order to maximize profits. As all innovative goods are symmetric and as there is free entry into R&D, profits $\pi(j)$ must be the same for all innovators and equal to F, the fixed costs of undertaking R&D, in equilibrium. This gives the free entry condition

$$\pi(j) = F \tag{3}$$

Let us define $\hat{\theta} \equiv \frac{\hat{y}}{\frac{Y}{L}}$, so that households of type $\theta < \hat{\theta}$ only consume a subset of the existing N innovative goods while households of type $\theta > \hat{\theta}$ consume all N innovative goods and some non-innovative goods. In the following, the case is considered in which θ has positive density $g(\theta)$ in the interval $\left[\underline{\theta};\bar{\theta}\right]$ and in which $\underline{\theta} < \hat{\theta} < \bar{\theta}$ holds, implying that there is a positive measure of households of both types (i.e. $\theta < \hat{\theta}$ holds for some of them and $\theta > \hat{\theta}$ for others).

As households differ with respect to income and the variety of innovative goods that they consume, not all innovators can sell to the same number of households in equilibrium. In order to guarantee equal profits for all innovators, the equilibrium price structure has to be such that innovators selling to more households sell at a lower price. As only richer households consume the more expensive goods, we can denote the price of the innovative good that is consumed by all households of type $\theta \geq \hat{\theta}$, but not by poorer households, by $p(\theta)$. Profits of the firm selling at price $p(\theta)$ to the number $L(1 - G(\theta))$ of households are then given by

$$\pi(\theta) = L(1 - G(\theta))\,(p(\theta) - c) \tag{4}$$

Setting these profits equal to the R&D costs F (equation (3)) allows to derive

$$p(\theta) = c + \frac{F}{L(1 - G(\theta))} \tag{5}$$

Prices therefore rise in the marginal production costs c and in the R&D costs F and fall in the market size $L(1 - G(\theta))$.

3.5 Solving for the equilibrium

The innovator that sells to all households of type $\theta \geq \hat{\theta}$ sets the limit price at which the household of type $\theta = \hat{\theta}$ is indifferent between purchasing the most expensive innovative good at price $p(\hat{\theta})$ and purchasing a non-innovative good at price $p_n = \Omega$, implying that $p(\hat{\theta}) = z_j(\hat{\theta}) = \Omega\frac{\partial f(C(\hat{\theta}))}{\partial C(\hat{\theta})}$ must hold[4]. Inserting this into equation (5) and taking into account that $C(\hat{\theta}) = N$, i.e. that the household of type $\hat{\theta}$ consumes one of each of the existing N innovative goods, the free entry condition can be written in the following way:

$$\left(\Omega\frac{\partial f(N)}{\partial N} - c\right) L(1 - G(\hat{\theta})) = F \tag{6}$$

[4] As mentioned above, the case is considered in which there is a positive density $g(\hat{\theta})$ of households of type $\theta = \hat{\theta}$.

When the limit price $p(\hat{\theta}) = \Omega\frac{\partial f(N)}{\partial N}$ increases due to an increase in the price (i.e. the production costs) Ω of non-innovative goods or due to a reduction in N (that implies an increased usefulness of the marginal innovative good relative to non-innovative goods as $\frac{\partial f(N)}{\partial N}$ falls in N), the innovator needs to sell to less individuals $L(1-G(\hat{\theta}))$ in order to break even and the equation is satisfied for a larger value of $\hat{\theta}$. When the size of the population L increases (taking total income and the limit price as given), the free entry condition is also satisfied for a larger value of $\hat{\theta}$ as the innovator can break even by selling one unit of the good to a smaller fraction of the population (i.e. to all households of type $\theta \geq \hat{\theta}$) if there are overall more households.

The budget constraint of a household of type $\theta < \hat{\theta}$ is given by

$$\theta\frac{Y}{L} = \int_{q=\underline{\theta}}^{\theta} p(q)dC(q) \tag{7}$$

where $dC(q)$ indicates the density of innovative goods sold at price $p(q)$. Differentiating with respect to θ gives $\frac{Y}{L} = p(\theta)dC(\theta)$. Solving for $dC(\theta)$ and integrating gives the variety of innovative goods consumed by the household as

$$C(\theta) = \frac{Y}{L}\int_{q=\underline{\theta}}^{\theta} \frac{1}{p(q)}dq \tag{8}$$

Inserting $p(q)$ from equation (5), the equilibrium number of innovative goods can be derived as

$$N = C(\hat{\theta}) = \frac{Y}{L}\int_{q=\underline{\theta}}^{\hat{\theta}} \frac{1}{p(q)}dq = \frac{Y}{L}\int_{q=\underline{\theta}}^{\hat{\theta}} \frac{L(1-G(q))}{cL(1-G(q))+F}dq \tag{9}$$

In Figure 1 this equation is plotted as curve "BC" ("budget constraint") and together with equation (6) (curve "FE' ("free entry")) it determines the equilibrium values of N and $\hat{\theta}$.

Proposition 1. ***a)****: Suppose that θ is distributed with positive support in the range $\underline{\theta} \leq \theta \leq \bar{\theta}$, where $\underline{\theta}$ ($0 \leq \underline{\theta} < 1$) is a sufficiently small and $\bar{\theta} > 1$ is a sufficiently large exogenous parameter. A unique equilibrium then always exists if $\left(\Omega\left(\lim_{N\to\infty}\frac{\partial f(N)}{\partial N}\right) - c\right)L > F$ holds.*

b)*: A progressive transfer among households of type $\theta < \hat{\theta}$ reduces N, while a progressive transfer from a household of type $\theta > \hat{\theta}$ to a household of type $\theta \leq \hat{\theta}$ increases N. Transfers between households of types $\theta > \hat{\theta}$ do not affect N.*

c)*: $\hat{\theta}$ increases in Ω and in L and decreases in Y (when $\frac{\partial f(C)}{\partial C} = 1\ \forall C$, $\hat{\theta}$ is independent of Y). Therefore, a progressive transfer between two randomly drawn households is the less likely to increase N (and the more likely to reduce N) the larger L and Ω are and the lower Y is (when $\frac{\partial f(C)}{\partial C} = 1\ \forall C$, the effects that the transfer has on N are independent of*

Y). N depends positively on Y and Ω and negatively on c and F (the effects of L on N and of F and c on $\hat{\theta}$ can be either positive or negative).

Proof. See appendix A1 □

Part b) of this proposition contains the main result of Föllmi and Zweimüller (2016) that a reduction in inequality can either increase or decrease innovation (they study a dynamic setup and analyze the effect of inequality on growth, assuming that $\frac{\partial f(C)}{\partial C} = 1\forall$ holds). This part implies that a progressive transfer from a randomly drawn household from the top x percent of the population to a randomly drawn household from the bottom $1 - x$ percent of the population is less likely to reduce N the smaller x is and that it (weakly) reduces N when x is sufficiently small[5]. This implies that reducing inequality by reducing the income share of the x percent richest households is less likely to reduce N the smaller x is and (weakly) increases N when x is sufficiently small. Part c) provides the basis for the main theoretical contribution of this paper. When inequality is measured using a standard Lorenz consistent measure like the Gini coefficient or the coefficient of variation, it falls when a progressive transfer takes place, but does not depend on the size of the population L or on total income Y when the distribution $G(\theta)$ of household types remains unchanged. Because of that, Proposition 1 implies the following:

Corollary 1. ***a)**:Suppose that the conditions from Proposition 1a hold. Reducing inequality is then the more likely to decrease (and the less likely to increase) the number of innovations N the larger the size of the population L, the larger the limit price parameter Ω, and the smaller total income Y is (when $\frac{\partial f(C)}{\partial C} = 1\ \forall C$, the effect of inequality on N does not depend on Y).*

***b)**: Reducing the income share of the richest x percent of the households is the less likely to reduce N the smaller x is and (weakly) increases N when x is sufficiently small.*

The intuition behind this result is the following: Whether redistributing income from a rich to a poor household encourages innovation depends on the consumption pattern of both types of households. When both only purchase innovative goods, a progressive transfer reduces the total number of innovations by shifting demand towards the less

[5]This can be derived in the following way: Let us define the household type at the x'th percentile of the distribution by θ_x, so that $G(\theta_x) = 1 - x$ holds. When x is small so that $\theta_x > \hat{\theta}$ holds, such a progressive transfer increases N when it is received by a household of type $\theta < \hat{\theta}$ and leaves N unchanged when it is received by a household of type $\theta > \hat{\theta}$. When x is so large that $\theta_x < \hat{\theta}$ holds, the transfer increases N when it originates from a household of type $\theta > \hat{\theta}$, but reduces N when it originates from a household the type of which lies in the interval between θ_x and $\hat{\theta}$. Such a progressive transfer is therefore less likely to reduce (and in the case where $\theta_x < \hat{\theta}$ also more likely to increase) N the smaller $\hat{\theta} - \theta_x$ is, i.e. the smaller x is.

exclusive goods that are sold at lower markups[6]. Due to this price effect a reduction in inequality can therefore reduce innovation.

When income is, however, redistributed from a rich household who already consumes one of each of the invented innovative goods and spends additional income on non-innovative goods to a poor household who only consumes innovative goods, this increases innovation through a market size effect: As the price setting power of firms supplying innovative goods is restricted due to the presence of substitutable non-innovative goods, inventing an innovative good is only worthwhile if sufficiently many households are rich enough to purchase it. A reduction in inequality therefore increases innovation when it leads to an increase of the number of goods for which there is "mass consumption" and when it reduces the consumption of a minority of rich households who's (incremental) demand has no effect on the incentives to innovate. Whether a reduction in inequality increases or reduces innovation therefore depends on the consumption pattern of the affected households as this determines whether the price or the market size effect dominates.

When the size of the population increases, innovators can break even when they sell to a lower fraction of the population, implying that a larger fraction of the households (i.e. all of type $\theta < \hat{\theta}$) spend all their income on innovative goods in equilibrium. This implies that the price effect becomes the dominating one for a larger fraction of households, making it less likely that a reduction of inequality stimulates innovation (and more likely that it reduces innovation). When the price Ω of non-innovative goods increases, innovators can charge a larger limit price and can break even if they sell to a smaller fraction of the population. It is then also less likely that a reduction in inequality stimulates innovation as the market size effect becomes less relevant.

When total income increases, the value of the marginal innovation decreases (when $\frac{\partial^2 f(C)}{\partial C^2} < 0$ holds) and the limit price falls. This implies that innovators need to sell to a larger fraction of the population in order to break even. As a larger fraction of the population then spends some income on non-innovative goods in equilibrium, the market size effect becomes more relevant and it becomes more likely that a reduction in inequality is good (and less likely that it is bad) for innovation. When $\frac{\partial f(C)}{\partial C} = 1 \; \forall C$ holds so that the value of the marginal innovation is independent of the number C of innovative goods consumed, the limit price is independent of total income and so is the effect of inequality on innovation.

It should be noted that the total number of innovations N is not simply a function of

[6] When a progressive transfer occurs, it reduces the number of innovative goods consumed by the rich household less than it increases the number of these goods consumed by the poor household (as the latter purchases goods at lower prices). This increases the overall production costs of firms supplying innovative goods, implying that fewer firms find it profitable to undertake R&D.

the number of households with an income larger than a certain threshold in this setup, but that it depends on the whole distribution of income. The reason for this is that this distribution endogenously determines the equilibrium price structure of innovative goods, implying that the willingness to pay of a household for a new good not only depends on this household's income, but on the prices of all other goods and therefore on the incomes of all other households.

3.6 International Context

The previous analysis considered a closed economy. When innovations are public goods, i.e. when the fixed costs of undertaking R&D have to be born only once, it is, however, likely that the incentives to innovate in one country do not only depend on the product demand and the extent of inequality within this country, but also on the demand for the innovations coming from other countries. In order to empirically test the predictions of the model, one therefore needs to overcome the difficulties involved in defining the market that is relevant for innovators. A further problem is that innovations might differ considerably across countries (in terms of usefulness for consumers, cost saving potential, or R&D costs), making it difficult to isolate the effect of inequality on innovation by simply comparing innovation rates in different countries. In order to get around these problems, the model is extended to an international context and a particular feature of the patent system is used: as no globally valid patents are granted, inventors only obtain patent protection in the countries in which they apply for it. As there are fixed costs associated with country-specific patent applications, we can therefore observe variation in patenting for given innovations across well-defined markets. The model is then used to show that this variation in patenting across countries should depend in the same qualitative way on the level of inequality within countries as innovation does in the simple closed economy setting. This result motivates our empirical analysis in which we study international patent applications in order to get around the problems that are associated with directly studying innovation outcomes. In the following, the model setup is described. There are M countries indexed by s. The size of the population of country s is given by L_s and the total endowment with efficiency units of labor by Y_s. Household types θ_s in country s are distributed according to the density function $g_s(\theta_s)$ with positive support on the interval $\underline{\theta}_s < \theta_s \leq \bar{\theta}_s$. Preferences are given by equation 1 and are assumed to be the same in all countries. The cost parameters in country s are given by c_s and Ω_s.

There are two types of fixed costs associated with innovation: In order to invent a generic innovative good, the invention costs $R > 0$ in terms of labor have to be incurred.

These costs are assumed to be the same in all countries[7] and there is free entry into the invention business. In order to adapt a generic invention to the particular conditions of a country s and in order to obtain patent protection on this good there, the additional fixed costs $F_s > 0$ in terms of labor from country s have to be incurred. These fixed adoption costs consists of direkt patent application costs, translation costs, or other costs that can for example arise when a good needs to be modified in order to be compatible with local standards. Therefore, F_s can be country-specific.

If an inventor does not patent an invention in country s, there is free entry into the imitation business, implying that a competitive fringe of firms can supply imitates of the good at marginal cost once the fixed costs of adoption, F_s, have been sunk. Anticipating this, no firm without patent protection finds it worthwhile to pay these fixed costs, as doing so would not allow to earn any positive profits ex post. When an inventor obtains patent protection in country s, other firms are prevented from supplying the same good there and transferring the technology might become worthwhile. This implies that only countries in which a particular innovative good j is protected by a patent can have access to this good in equilibrium.

An alternative scenario without technology adoption costs in which an invention can be freely adopted and supplied at marginal cost when it is not protected by a patent is analyzed in **Appendix A3**. When all (or most) households are rich enough to purchase all non-patented goods, the main predictions of the model remain the same in this case[8]. Whether or not patent applications are associated with a transfer of technology or merely an extraction of monopoly rents therefore does not seem to matter much for the effect of inequality on patenting.

It is assumed that trade costs are sufficiently large to make it unprofitable to ship an innovative good to a country in which it is not profitable to patent it. Moreover, it is assumed that parallel trade is prohibited for patented innovations, allowing patent holders to charge different prices in different countries. These assumptions that limit the role of trade are made in order to keep the analysis simple and tractable. As it is is likely that trade relations between countries affect patenting decisions in the real world, we nevertheless control for trade flows in our empirical analysis and show that including these does not affect our (qualitative) results.

When the measure N_s of innovative goods are invented in country s, the global measure of inventions (the "world technological frontier") is given by $N = \sum_{s=1}^{M} N_s$. In country s, the subset $V_s \leq N$ of all available inventions are patented.

[7]This assumption is made in order to obtain the simplest possible equilibrium in which no goods need to be traded across countries.

[8]A welfare analysis can, however, lead to different conclusions in both cases

3.6.1 Equilibrium

As adopting already invented goods is cheaper than inventing and adopting new ones and as all innovative goods are symmetric, inventing can only be profitable if there is at least one "frontier" country in which all the globally available inventions are adopted, i.e. in which $V_s = N$ holds. In other "follower" countries, only a fraction of the globally available innovations might, however, be adopted, i.e. $V_s < N$ might hold. Within this setting, whether a country is a frontier or a follower country is endogenously determined and depends on the relative profitability of its market (which is a function of different parameters, in particular Y_s and L_s).

In a follower country in which $V_s < N$ holds, inventors must be indifferent between applying and not applying for patent protection. This is the case if their ex post profits from selling their good in such a country s are equal to the fixed costs F_s of adoption. As these fixed costs consist of labor from country s and as, due to the lack of parallel trade, the optimal prices charged by patent holders only depend on local demand conditions, the equilibrium in a country in which $V_s < N$ holds is similar to the one in the basic closed economy studied above. The only difference is that the endogenous number of innovations N is now replaced by the endogenous number of (international) patent applications V_s and that the fixed costs F of innovating are replaced by the fixed costs F_s of adoption (note that a situation in which more inventions are globally available than are adopted in a certain country is similar to one of "free entry" into adoption). Because of this, V_s depends in the same way on the parameters in country s as in the closed economy and the results of Proposition 1 can be directly applied with N replaced by V_s. When innovations are randomly adopted from the world technology frontier, the probability that a given invention is patented in country s is then given by $\frac{V_s}{N}$.

If there is only one "frontier" country in which $V_s = N$ holds and in which sufficient profits can be earned in excess of the adoption costs in order to make R&D worthwhile, the world technology frontier N is pinned down by the parameters of this country like in the closed economy model studied above. When some of the innovations patented in this frontier country are also patented in other "follower" countries in which $V_s < N$ holds, this does not affect the profitability of R&D as no net profits are derived from obtaining patent protection in those follower countries. If there are $H > 1$ frontier countries in which $V_s = N$ holds, the following free entry condition needs to hold in each of these countries:

$$\sum_{s=1}^{H} (\pi_s - F_s) = R$$

This implies that the world technological frontier N is a function of the parameters in all these H frontier countries. As R&D costs R are assumed to be the same in all countries, firms are indifferent about where to undertake the R&D. Consequently, the equilibrium can be considered in which each frontier country undertakes the level of R&D that coincides with the "net profit income" $N_s(\pi_s - F_s)$ derived in this country and in which no profit income needs to be transferred across countries (note that profits net of all fixed costs are equal to zero due to free entry into R&D).

While innovations are symmetric in the model, they are not symmetric in the real world, where some are clearly more likely to be adopted and patented abroad than others. This might be due to technological features of particular innovations, their value, or different (trade) relations between the country of invention (or first patent application) and the country of adoption. While the model would become much more complicated to solve if it would explicitly allow for heterogeneous innovations or trade, some heterogeneity can nevertheless be introduced in a simple way: as inventors are in equilibrium indifferent about whether or not to adopt their invention in a country in which $V_s < N$ holds, we can assume that the probability that a particular invention is patented in such a country does not only depend on the relation between V_s and N, but also on other (not modelled) characteristics of the technology, characteristics of the country of invention (or first patent application) and characteristics of the (trade) relationship between the country of invention and the country of adoption. Formally, the probability Q_{jkl} of innovative good j invented (or first patented) in origin country $s = k$ to be patented (adopted) in destination country $s = l$ is given by

$$Q_{jkl} = h(V_l, N, z_{jkl})$$

where z_{jkl} is a vector of variables specific to the innovation, the origin and the destination country and where $\frac{\partial h(V_l,N,z_{jkl})}{\partial V_l} > 0$ and $\frac{\partial h(V_l,N,z_{jkl})}{\partial N} < 0$ holds. When a group i of innovations which all share the same characteristics z_{ikl} (including the country of origin) is considered, Q_{ikl} simply denotes the fraction of these innovations that gets patented in destination country l. In the following, Q_{ikl} is denoted as the "patent flows" from country k to country l occurring for technology i. As the relation between V_l and the parameters of the adopting country l is the same as in Proposition 1 (with V_l replaced by N), we can state the following proposition

Proposition 2. *Suppose that the conditions from Lemma 1a hold in a follower country l in which $V_l < N$ holds.*

a): *A reduction in inequality I_l in country l is the more likely to decrease (and the less likely to increase) patent flows Q_{ikl} from origin country k to destination country l for technology i the larger the population size L_l and the limit price parameter Ω_l are and the lower total income Y_l is.*

Formally, $\frac{\partial E\left(\frac{\partial Q_{jkl}}{\partial I_l}\right)}{\partial L_l} > 0$, $\frac{\partial E\left(\frac{\partial Q_{jkl}}{\partial I_l}\right)}{\partial \Omega_l} > 0$ *and* $\frac{\partial E\left(\frac{\partial Q_{jkl}}{\partial I_l}\right)}{\partial Y_l} < 0$ *hold (in the case where* $\frac{\partial f(C)}{\partial C} = 1 \; \forall C$, $\frac{\partial E\left(\frac{\partial Q_{jkl}}{\partial I_l}\right)}{\partial Y_l} = 0$*), where* $E\left(\frac{\partial Q_{jkl}}{\partial I_l}\right)$ *denotes the average (expected) change in* Q_{jkl} *over all possible regressive transfers (of given size) between household pairs due to which inequality might be increased.*

Moreover, Q_{ikl} *depends positively on* Y_l *and* Ω_l*, negatively on the cost parameters* c_l*,* F_l *and negatively on the world technological frontier* N *(the effect of* L_l *on* Q_{ikl} *is ambiguous).*

b): b): *Reducing the income share of the richest* x *percent of the households is less likely to reduce* EQ_{ikl} *the smaller* x *is and increases* EQ_{ikl} *when* x *is sufficiently small.*

It should be noted that the proposition can also be directly applied to the total number of innovations from technology i getting patented in l.

This proposition shows that inequality in country l affects patent applications in this country in the same qualitative way in which inequality affects innovation in the closed economy. Empirically studying how patent applications across countries depend on the level of inequality might therefore allow to make inferences about how inequality affects innovation.

3.7 Extensions

3.7.1 Non-innovative basic need goods

In the model above, poor households spend all their income on innovative goods and only rich households also purchase some non-innovative goods. In **Appendix A2**, an extension is analyzed in which there are some non-innovative "basic need goods" like food[9] which all households consume before they start consuming innovative goods. When all (or most) households are rich enough to purchase all of these basic need goods, the qualitative results of the analysis stay the same.

3.7.2 Limited strength of patent protection

This section analyzes how the effect of inequality on innovation depends on the strength of patent protection, modeling the latter as either patent breadth or a varying probability of enforcement.

Suppose that once a technology is patented and adopted in a country, imitators can still enter the market as long as their imitates are sufficiently worse compared to the patented innovation. The **breadth of patent protection** then determines how much

[9]According to many empirical studies, the budget share on food falls in household income. This is called "Engel´s law" and referrs to Engel (1857)

worse imitates have to be compared to a patented innovaton in order not to infringe on the patent[10]. In this case, reducing the breadth of patent protection has the same effect as reducing the limit price parameter Ω as it restricts the price setting power of innovators. A smaller patent breadth therefore leads to a reduction in $\hat{\theta}$ and makes it less likely that inequality is good for innovation and more likely that it is bad for innovation.

Suppose now that instead of the breadth of patent protection the **probability of patent enforcement** is varied: Let us assume that inventors who have already paid the fixed costs F_l to patent their innovation in country l only obtain patent protection there with probability Δ_l and are imitated with probability $1 - \Delta_l$. In the case of imitation, there is Bertrand competition and the price falls to the marginal costs c_l, implying zero profits. Studying the simple case in which $\frac{\partial f(C)}{\partial C} = 1 \ \forall C$, the free entry condition in country l is then given by

$$\Delta_l \left(\Omega_l - c_l\right) L_l(1 - G(\hat{\theta}_l)) = F_l$$

Consequently, $\hat{\theta}_l$ depends positively on Δ_l. This is because an increase in Δ_l allows a firm that sells at the limit price Ω_l to still break even if it sells to fewer households when the patent is enforced.

As consumers prefer to purchase the cheap innovative goods on which patents are not enforced, they only consume patented innovative goods when they are rich enough to purchase one of each of the $(1 - \Delta_l) V_l$ competitively supplied goods. Let us only consider the simple case in which all households are rich enough to purchase some patent protected goods[11]. A regressive transfer among households of type $\theta_{il} < \hat{\theta}_l$ then still increases V_l, while a regressive transfer from a household of type $\theta_{il} < \hat{\theta}_l$ to a household of type $\theta_{il} > \hat{\theta}_l$ still reduces V_l (the proof resembles that provided in Appendix A2). Because of that, an increase in patent enforcment Δ_l in country l makes it more likely that inequality is good and less likely that it is bad for patent applications in this country (as it increases $\hat{\theta}_l$).

Summing up, the following proposition holds:

[10]This could be modelled in the following way: suppose that consumers consider an imitated good to be of equal value as a non-innovative good (independently of how many innovative goods they consume). Then, the breadth of patent protection would determine a lower bound P on production costs of imitated goods below which imitators would infringe on the patents of innovators. Consequently, the new limit price for innovators would be given by $P\frac{\partial f(N)}{\partial N}$ if $B < \Omega$ holds.

[11]This is the case if $\underline{\theta}_l \frac{Y_l}{L_l} > c_l \left(1 - \Delta_l\right) V_l$ holds. Given that parameters are such that an equilibrium with $V_l > 0$ exists when $\Delta_l = 1$ holds (sufficient conditions for that are stated in Lemma 1a), continuity of the BC and the FE curve in Δ_l imply that this inequality is satisfied in equilibrium when Δ_l is sufficiently large (the BC curve is given by $V_l = \frac{Y_l}{L_l} \int_{q=\frac{c_l(1-\Delta_l)V_l L_l}{Y_l}}^{\hat{\theta}_l} \frac{L_l(1-G(q))}{c_l \Delta_l L_l (1-G(q)) + F} dq$, where $\frac{c_l(1-\Delta_l)V_l L_l}{Y_l}$ indicates the level of θ_l above which households start consuming patent protected goods, and where the integrand is derived from the new free entry condition $p(\theta_l) = c_l + \frac{F_l}{\Delta_l(1-G(\theta_l))L_l}$ (that replaces equation 5)).

Proposition 3. *In both cases analyzed in this section, strengthening patent protection in a destination country l makes it more likely that inequality in this country increases (and less likely that it decreases) patent applications there.*

Formally, $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial P_l} > 0$ *holds when P_l indicates the strength of patent protection in country l.*

Kiedaisch (2016) studies an endogenous growth version of this model. Looking at the case of two income groups (that are both rich enough to purchase all non-patented goods), he also finds that reducing patent breadth or increasing the probability of patent expiration makes it less likely that inequality is good for growth. Taking transitional dynamics into account, he, moreover, shows that strengthening patent protection might not always increase growth when it leads to changes in the level of inequality. As we control for the level of inequality in our empirical analysis, we do not consider such additional effects here.

4 Empirical Analysis

In the following empirical part, we want to confront the theoretical predictions as stated in Proposition 2 and Proposition 3 with data. We therefore examine the comparative statics results implied by the theoretical discussion in the previous section, at the level of a single technology i initially patented in country k and patent protection extended to country l at time t. We therefore consider a sample of follower countries in which a positive amount of patented innovations from abroad is observed. Thus we only consider technologies for which a subsequent patent application is filed in at least one other country.

We then look at the variation in patent flows given characteristics of the origin and destination countries and the technologies i. As we consider panel-data, some combinations ikl surface repeatedly over time. We consider such repeated filings as indicative of a higher probability that the technology truly enters the foreign market.

4.1 Design and Estimation

We parametrize the conditional expectation of the measure of international patent-flows Q_{iklt} given explanatory variables C_{iklt} taking on value c_{iklt}

$$\mathbb{E}(Q_{iklt}|c_{iklt} \in C_{iklt}) = \exp(\eta_{iklt}) \tag{10}$$

with the parameter

$$\eta_{iklt} = c_{iklt}^T \alpha \tag{11}$$

a linear index in the elements of vector c_{iklt} and conformable coefficient vector α. We parametrize this index as

$$\eta_{iklt} = \iota^T \underline{\delta}_{iklt} + z_{iklt}^T \gamma + x_{lt}^T \beta \tag{12}$$

where x_{lt} contains realizations of the theoretically motivated variables of interest with conformable coefficient vector $\beta \in \alpha$, z_{iklt} contains further control variables (with according parameters collected in $\gamma \in \alpha$), and $\iota^T \underline{\delta}_{iklt} = \delta_i + \delta_k + \delta_l + \delta_t$ represent effects specific to the according dimension of variation of patent-flows Q_{iklt}. As the continuum of possible technologies is infinite, we consider unique technologies i as the incidental dimension of the data, whereas the other dimensions of variation are considered as fixed strata of origin, destination, and time. We will therefore assume that $\delta_i = \delta_o + \kappa_i$, where δ_o is a fixed effect over a finite set of broad economic sectors, and κ_i is innovation-specific. We will assume that $\kappa_i = 0$ or if non-zero it is a random variable exogenous with respect to the elements of (z_{iklt}, x_{lt}) and with conditional expectation set to zero. In the latter case we can ensure valid inference by according clustering of standard-errors.[12] In a robustness check below, we will allow for $\delta_i \neq \delta_o$ and δ_i correlated with elements of (z_{iklt}, x_{lt}) via κ_i.

4.1.1 Theoretical Relationship

As measures of a country's population and income we use L_{lt} =log(POP$_{lt}$) (the logarithm of a follower country's population) and Y_{lt} =log(GDP$_{lt}$) (the logarithm of a follower country's GDP), respectively. For a Lorentz-consistent measure of inequality I_{lt}, the empirical relationship implied by the result as stated Proposition 2 is

$$x_{lt}^T \beta = \beta_1 I_{lt} + \beta_2 I_{lt} \log(\text{POP}_{lt}) + \beta_3 I_{lt} \log(\text{GDP}_{lt}), \tag{13}$$

where we would expect that $\beta_2 > 0$, $\beta_3 < 0$, and $\beta_1 \lesseqgtr 0$.

To consider the theoretical prediction of Proposition 3 we add an additional interaction term in a measure of patent protection (B_l in the proposition) and inequality. For the coefficient on this interaction, β_4, we would in addition expect that $\beta_4 > 0$.

In order to study how inequality affects patent flows according to (13), we can calculate

[12] Or as we will apply in the light of a generated regressor in an additional robustness exercise below, we may estimate a bootstrap-confidence interval by a resampling with replacement form over varying instances of i (o) with a fixed klt.

the decision boundary of a positive versus a negative partial effect of inequality, that is

$$d(I_{lt}) = \{(POP_{lt}, GDP_{lt}) : \partial x_{lt}^T \beta / \partial I_{lt} = 0\}, \tag{14}$$

i.e. all combinations of population and GDP for which the marginal effect of inequality is exactly zero. If we want to include another variable – as for exploring Proposition 3 – we accordingly obtain a decision surface with coordinates population, GDP, and amount of patent protection.

4.1.2 Estimation

The exponential conditional mean model $\exp(\eta_{iklt})$ motivates estimation of a generalized linear model (GLM) as introduced in Nelder and Wedderburn (1972), Wedderburn (1974), and McCullagh and Nelder (1983). Considering one-parameter exponential families for the conditional distribution of the dependent variable given explanatory variables, the conditional mean of interest can be estimated consistently even if the assumed distributional model is mis-specified (see Gourieroux et. al; 1984). Conditional distributions within this class that are consistent with the support of the conditional mean of patent-flows are the Poisson-distribution, the one-parameter Gamma-distribution, and the negative binomial distribution using the quadratic mean parametrization (NB2). As a further approach motivated by the exponential relationship in (12) we may log-transform the dependent variable and apply linear least squares (or a gaussian GLM with an identity instead of the logarithmic link as is employed here). However, this is a model for a transformed relationship and as implied by Jensen's inequality we may interpret the resulting conditional expectation as lower bound on the conditional expectation for the original relationship.

To compare among a set of different estimated econometric models for the conditional expectation of interest we refer to in- and out-of-sample fit. To assess in-sample fit we look at the squared correlation among the true and the fitted values for each model. To measure out-of-sample fit – i.e. how well any fitted regression would perform when predicting a new data-point on the dependent variable from observed values of the explanatory variables – we compute the cross-validation-error.

4.2 Dependent Variable

To construct a measure of patent-flows we utilize information comprised in the PATSTAT-database.[13] We extract priority applications for the years 1980-2013 of granted patents of

[13] The PATSTAT database is published by the European Patent Office and contains world-wide data on patents. The analysis proposed in this paper is based upon its 2015 Autumn Edition.

invention as according to Article 4 of the Paris Convention of the Protection of Industrial Property 1883, that are followed by a subsequent filing at a different patent-office, i.e. any initial application that is being claimed as a priority elsewhere within 1980-2013. [14] [15]

A raw-patent-flow occurs if a priority-fling filed at one patent-authority is found within a subsequent filing at a different patenting-authority in a year.[16] We measure the timing of a patent-flow as the year when the priority is cited at the level of the receiving-office.[17] We consider subsequent filings as arising after the Paris Convention and the Patent Co-operation Treaty 1970 (PCT).[18] Both, the Paris Convention and the PCT experienced substantive amendments by the end of the year 1979 motivating our restriction of attention to periods since 1980. In our raw-data, a share of roughly 86.7% of all subsequent filings is due to Paris Convention direct filing, whereas the rest are PCT applications.[19]

Having obtained raw-patent-flows, we aggregate those for distinct combinations of technology i, origin k, destination l and year t. To measure distinct technologies i we refer to distinct PATSTAT DOCDB-families. Alternatively, we consider different family-definitions by means of a single-priority-family – where each distinct priority filing is considered as a distinct technology – and the extended patent-family definition INPADOC. We normalize the size of each technology to one such that in case a technology consists of several priorities each subsequent application of a single priority enters with a fractional weight. At the level of the subsequent filing one or two priorities may be referred to

[14]In PATSTAT such priorities can be identified by their absence from table `TLS204_APPLN_PRIOR` and `INTERNAT_APPLN_ID` in table `TLS201_APPLN` being 0, together with the information on granting status contained in table `TLS201_APPLN`. (Artificial or replenished observations are excluded)

[15]Due to data-availability for the explanatory variables introduced below, the most recent year included in the analysis is 2013.

[16]As an alternative measure of geographical location one might consider the nationality of the inventor(s) (cf. Eaton et. al. 2006). We decided to follow the view that the relevant market of protection is more likely associated with the location of the patent office.

[17]We consider this as in-line with the different timing structures between priority date and the date when a subsequent application faces at the level of a national office for Paris Convention direct-filing an PCT-filings (entering their national phase), respectively. Thus any considerations with respect to strategically exploiting the different timing-structures by the applicant to optimize provisional patent protection are then assumed to occur previous to the patent-flow facing in the data.

[18]In PATSTAT such observations are found by matching the content of the variable `PRIOR_APPLN_ID` in table `TLS204_APPLN_PRIOR` to the variable `APPLN_ID` for those observations previously identified as priority filings; their `APPN_KIND` is either “A" for Paris Convention subsequent filings or ‘W" for PCT. Note that this refers to PCT filings as the enter their national phase.

[19]As compared to direct filing, PCT filing provides the applicant a longer period of provisional patent protection and costs occur at a later point in time. PCT filing is generally more expensive than direct filing, but is usually more cost efficient when protection in several foreign countries shall be obtained simultaneously. Hence, large international companies might be more likely to file after the PCT, whereas SMEs might be more likely to follow the direct filing strategy of the Paris Convention. The Paris Convention has wider country-coverage than the PCT, and accordingly for some countries patent protection can only be obtained by direct filing at all. Moreover – which is also a widely used strategy in practice – international patent protection can also be applied for simultaneously after both international treaties.

as prior art, and in such a case a fractional flow from each priority is considered. Thus for several raw-patent-flows within a fixed combination of $iklt$ the dependent variable is defined as

$$Q_{iklt} = \sum_{g \in iklt} \mathbb{I}\{g \in iklt\} w^{\text{family}}_{g(ik)} * w^{\text{citing}}_{g(lt)} := Q^{w}_{iklt}$$

with $\sum_{g \in ik} w^{\text{family}}_{g(ik)} = 1$ and $\sum_{g \in lt} w^{\text{citing}}_{g(lt)} = 1$ and $\mathbb{I}$ is the indicator function. For a single-priority-family $w^{\text{family}}_{g(ik)} = 1$, whereas a DOCDB- or INPADOC-patent-family can consist of more than one priority-filing which can also be attached to several countries and therefore we choose

$$w^{\text{family}}_{g(ik)} = 1/\text{“No. of priorities the family consists of”}$$

for a priority that belongs to family i and was filed in k. Observe that for the single-priority-family the combination of i and k is necessarily fixed, and a technology cannot be distributed across several countries of origin. Analogously, for the weight attached to the destination in case more than one – in the data we observe at most two – priority is cited

$$w^{\text{citing}}_{g(lt)} = 1/\text{“No. of priorities cited”}$$

where $\sum_{g \in lt} w^{\text{citing}}_{g(lt)} = 1$ for given destination l and year t. We may motivate the chosen uniform weights as the maximum entropy weights that are least informative about the importance of each innovation in that bundle.[20] As an alternative strategy we also consider an unweighted definition of the dependent variable as $Q_{iklt} = \sum_{g \in iklt} \mathbb{I}\{g \in iklt\} := Q^{u}_{iklt}$. As we will see below, the Q^{w}_{iklt} has a smoother empirical distribution with less mass points as compared to Q^{u}_{iklt}.

4.2.1 Descriptive Statistics for Dependent Variable

Following the previous descriptions there are several ways of defining the dependent variable. To convince the critical reader that we may just consider one out of them, we briefly illustrate the high overlap in patterns as suggested by several different definitions (as ac-

[20] Thus at the level of the priority the following cases can occur: (a) one-priority family or (b) multiple priority family, and at the level of the receiving office for the subsequent filing either (c) one priority cited or (d) two priorities are cited can occur, yielding to 2 x 2 distinct constellations. For (a) with (c) this yields to $w^{\text{family}}_{g(ik)} = 1$ and $w^{\text{citing}}_{g(lt)} = 1$, for (a) with (d) $w^{\text{family}}_{g(ik)} = 1$ and $w^{\text{citing}}_{g(lt)} = 1/2$, for (b) with (c) for e.g. a family consisting of 3 priorities $w^{\text{family}}_{g(ik)} = 1/3$ and $w^{\text{citing}}_{g(lt)} = 1$, and for (b) with (d) again for the example with 3 priorities $w^{\text{family}}_{g(ik)} = 1/3$ and $w^{\text{citing}}_{g(lt)} = 1/2$. Then for all distinct $iklt$ observed we sum-up over all instance s for a fixed combination $iklt$, e.g. for a given family i consisting of 4 priorities filed at the same origin k out of which 3 are cited by a PCT-application and 2 are cited by a Paris-Convention filing at the same destination l in the same year t and the subsequent applications do not cite any other priorities outside of the family, then $Q_{iklt} = 3/4 + 1/2 = 1.25$.

cording to three distinct family definitions and whether a weighted or and unweighted count is considered). Moreover, we vary the set of subsequent filings considered: (1) Paris-Convention and PCT subsequent filings, (2) only Paris-Convention filings, and (3) whether to include regional subsequent-applications directly except for calculating the weights of the weighted count.

Figure 2 highlights patent-flows as according to the resulting $3 \times 2 \times 3$ distinct definitions of the dependent variable, each at the aggregate level of distinct NACE2 Divisions of manufacturing, of countries of origin and destination, and time as surfacing in the data. The stars plots shown there visualize the relative size of each category (NACE2 Divisions, countries of origin and destination, years) with respect to the 18 distinct patent-flow measures as compared to all other categories. Different measures of the dependent variables are represented by different colors. What is highlighted there is a strong overlap in patterns, no matter which type of family definition is considered, no matter if PCT and Paris-Convention filings are considered together or whether or not regional applications are excluded. Without loss of generality, we will focus on non-regional subsequent applications according to the Paris Convention and the PCT in the following. These are represented by the radii of the variables A and B (weighted and unweighted count for the single-priority-family), G and D (weighted and unweighted count for the DOCDB-family), and M and N (weighted and unweighted count for the INPADOC-family).

Figure 3 plots the relative frequencies of the selected variables (in logs), and Table 1 supplements this information by a summary of descriptive statistics. The numbers of observations on data on the dependent variable is approximately 3.7 mio observations for each definition. The reported first and second moments show that relative to a Poisson-distributed random variable, the unweighted counts are over-dispersed whereas the fractional counts are under-dispersed. The over-dispersion of the unweighted count may be accounted to many entries equal to one and two in the data whereas the fractional counting smooths the mass-points a bit.

To motivate our later employed empirical specification of the fixed effects, we present an ANOVA for the log-transformed measures of patent-flows in Table 2. As indicated by the reported mean-squares, variation over different divisions of manufacturing as represented by δ_o above explains most of total variation, and variation over countries of destination (δ_l) second most. Variation over origins and time (δ_k and δ_t) is less pronounced. Moreover, for the weighted count variation over origins has a higher mean-square as variation over time, which holds for the weighted count vice versa.

We conclude that the patterns of variation in the dependent variable for different family-definitions are highly similar. In the sequel, we will therefore concentrate at the DOCDB-family as the entity defining distinct innovations, and will retain the other definitions for robustness checks.[21]

[Figure 2, Figure 3, Table 1, Table 2]

4.3 Explanatory Variables

4.3.1 Inequaltiy

A measure of inequality I_{lt} with good data availability is the Gini coefficient measuring the deviation of a country's income distribution from perfect equality. It takes on values on $[0, 1]$ with a value of 1 implying that one member of the population gets all the income whereas all others get none. A very rich data-set qualitatively oriented towards the gold standard of the Luxembourg Income Study on post- and pre-tax Gini coefficients (GINInet_{lt} and GINImarket_{lt} below) is provided by the Standardized World Income and Inequality data-base (Solt, 2016; SWIID).[22] The most recent year covered by the SWIID is 2013.

As a different measure of inequality we use top-income shares. The World Wealth and Income Database (WID; Piketty et. al) provides rich time-series-country-data on income and wealth statistics predominantly constructed from tax records. From this data-base we use information on top-10%, top-5%, and top-1% income shares ($\text{TOP10\%-share}_{lt}$, $\text{TOP5\%-share}_{lt}$, and $\text{TOP1\%-share}_{lt}$ as referred to below). As compared to the Gini coefficient, top-income shares immediately highlight whether inequality is attributable to a small rich class. Country-year-coverage for the WID is worse than for the SWIID, and especially poor for the years from 2013 onwards. As according to the latest period covered by the SWIID, we retain the period 2013 for top-income shares but drop future periods.

In Figure 4 we assess the ten distinct marginal relationships that exist among these five inequality measures whenever they are simultaneously observed for a destination l and year t surfacing in our data. The top-income-shares are almost linearly related when considering the three binary relations, and a similar pattern holds for GINInet_{lt} and $\text{TOP1\%-share}_{lt}$. Between GINInet_{lt} and $\text{TOP5\%-share}_{lt}$, and GINInet_{lt} and TOP10%-share_{lt} share there is an inverse U-shaped relationship between inequality post-taxes and high incomes. Qualitatively this tells us that high levels of inequality post-taxes are most likely accompanied by high top-1% incomes and comparably little mass below at top-10% and top-5% quantiles of the income distribution. Moreover, the observed pattern for the

[21] Further results for the analysis presented below using the other definitions of the level of the innovation i for defining the dependent variable are available from the authors upon request and will be made available in the Online Appendix to this paper.

[22] As the SWIID-data were constructed by multiple imputation, point measures of pre- and post-tax Gini coefficients are constructed by averages over all MI-estimates as reported by Solt.

Gini coefficients pre- and post taxes (GINInet_{lt} vs. GINImar_{lt} in the figure) indicate that for high levels of inequality (approx. > 0.45) there is almost a one-to-one relationship between the Gini coefficients pre- and post-taxes, whereas for low levels of inequality previous to taxes, the reduction in inequality at the level post-taxes is much larger in order of magnitude. Qualitatively, this means that countries in our sample that have low levels of inequality previous to taxes redistribute income more equally post-taxes as compared to countries with a high level of inequality among market incomes initially.

[Figure 4]

4.3.2 Population and GDP

Measures of population and real GDP are obtained from the World Bank's World Development Indicators (WDI).[23] As both variables have strictly positive support, we log-transform them. We will refer to the log of the destination's population and GDP at time t by $\log(\text{POP}_{lt})$ and $\log(\text{GDP}_{lt})$ as introduced above, and use $\log(\text{POP}_{kt})$ and $\log(\text{GDP}_{kt})$ when the origin-country is referred to, respectively. We consider the latter two variables as controls for characteristics of the origin-country of the patent-flow, interpretable as measures of economic size.

In Appendix B we provide some further results on pair-wise correlations among inequality measures and GDP and population in our data, respectively.[24]

4.3.3 Measures of Patent Value, Costs, and Strength of Patent Protection

There is common sense in the empirical literature that the number of citations a given patent, or a bundle thereof respectively, receives mirrors the economic value of inventions (see Trajtenberg, 1990; Harhoff et. al. 1999; Harhoff et al., 2003; Hall et al., 2005; Abrams and Akcigit; 2013). The study of Trajtenberg (1990) suggests that the social value of an invention is positively correlated with the incidence of subsequent citations. The results presented in Harhoff et. al. (2003) suggest that the number of references to the patent literature as well as the citations a patent receives are positively related to its value. Hall et. al (2005) consider how the market value of the firm as the holder of the patent responds positively to patent citations. More recently, Abrams and Akcigit

[23] The WDI-series references are SP.POP.TOTL (population) and NY.GDP.MKTP.CN.

[24] As indicated above, we confirm that there are positive correlations among all inequality measures. Moreover, the correlation of the share of the top-1% incomes with, population, GDP, and GDP per capita, respectively, are not estimated significantly different from zero, otherwise there is a positive correlation among top-5% and top-10% incomes with population and GDP, whereas there is a negative correlation between Gini coefficients with GDP per capita. This means, that for our data rich countries are likely to have a low level of inequality in terms of Gini coefficients, but a comparably high top-5%/top-10% income share.

(2013) have pointed at an inverted U-shaped relationship between the value of a patent and a patent's lifetime citations, though a positive correlation between lifetime citations and patent-value is confirmed.[25]

To construct a measure of the value of a technology, we construct a set of five variables measuring the spreading-out of subsequent applications at different points in time and over space: (1) A variable counting the number of a technology's previous subsequent applications up to time t, (2) another variable counting the number of contemporaneous applications of a technology at time t, (3) a variable using (1) normalized by the technology's age, (4) a variable counting the number of different destination countries where a technology has already been filed at in t, and finally (5) a variable counting the number of distinct countries where there is a subsequent application for the technology simultaneously in a year t. Technologies are therefore measured by any unique priority in case the single-priority-family is applied, or at the aggregate level of the DOCDB- or INPADOC-patent-family as recorded in PATSTAT, respectively.[26] In either case, all of the five variables are highly positively correlated, and after log-transforming them, we extract the first principal component which accounts for roughly 80% of the total variance and use it as a measure of patent-value, PATVAL_{it}. As is favourable for interpretability, all of the five variables show a positive loading on the first PC. Further information and estimation output is provided in the appendix.

As a further potential correlate of a technology's value in terms of the human capital equivalent required to produce the invention, we consider its number of inventors. For the single-priority-family definition of a technology this information is contained in PATSTAT as the number of inventors recorded with the priority filing. For the aggregate family-definitions we use the sum of inventors behind the bundle of unique priorities cited with a subsequent application. As this variable has strictly positive support, we apply a log-transformation that we will refer to by $\log(\text{NOINVENTORS}_i)$.[27]

To take account of the heterogeneous structure of cost associated with making a technology within different branches of the manufacturing industry, we include a set of indicator variables δ_o for the NACE2 Division within which the patent-flow occurs (cf. above), and a set of indicator variables δ_l for the destination of the patent-flow accounting for heterogeneous fee-structures among patent-offices. Since a major cost factor of foreign patent-extensions are due to translation cost, we include an indicator variable from the CEPII of whether the origin and the destination of the patent-flow share the same official

[25]They attribute this non-monotonicity to filing purposes that are strategic versus productive, where only in the latter case a positive correlation of citations and patent-value occurs.

[26]For multi-priority families with different filing years, the average age is considered as the age of family, i.e. innovaiton.

[27]Raw-observations with inconsistent inventor information were dropped from our analysis.

language (COMLANG_{kl}).

As a measure of patent rights in the foreign market where protection of the innovation is desired, we employ the index suggested in Ginarte and Park (1997), and updated by Park (2008). The index measures the extent of coverage of patent protection provided by a country's legal system, provisions of loss protection, enforcement mechanisms, the duration of protection, and membership in international patent agreements. Allred and Park (2007) find a positive relationship between strong IPR protection and foreign patenting. Aghion et. al. (2013) find a positive relationship between R&D intensity and patent protection as measured by the Ginarte-Park-Index. The Ginarte-Park-Index takes on values between zero and five, where higher values indicate stronger IPR protection, and its most current version the time span 1960-2010 with a measurement period of every five years. To construct data on an annual basis for the period 1980-2013, we firstly linearly interpolate each country-specific time-series on an annual grid, and then apply country-wise local linear regression to obtain predictions for the years 2011 until 2013. We will refer to this interpolated and extended variable as GP-INDEX_{lt} below.

4.3.4 Inventive and Imitative Capacity of Destination Country

One motivation of patenting abroad is the aim to protect one's innovation on the local market from imported imitations (Eaton and Kortum, 1996; Grupp and Schmoch, 1999; Peeters and van Pottelsberghe, 2006).

As a measure of the capacity of a destination country l to imitate the set of innovations it imports from to the origin k at time t we construct a technological similarity index TECHSIM_{klt} in the spirit of Jaffe (1986). The variable measures the proximity of two countries' technological profiles of exported patents in a year t, and is constructed by calculating the share of exported patents in each NACE2-Division for each country to then obtain the inner product of the two vectors as associated with a pair of countries kl. This variable will equal 1 for countries with exactly the same distribution of patents across classes, and 0 for countries with no patents in the same classes.

As another proxy of a destination country's inventive or imitative capacity we employ a measure of scientific output per capita based on raw-data extracted from the WDI.[28] To construct this measure, we add one the number one to each entry for the number of scientific and technical journal articles in the raw-data and then divide by the country's population. Since the resulting variable has strictly positive support, we apply the log-transformation. We will refer to this variable below by $\log(\text{ARTpc}_{lt})$.

[28]The original variables from the WDI are IP.JRN.ARTC.SC and SP.POP.TOTL (as used above).

4.3.5 Trade and Distance

Slama (1981), Park (2003), Eaton et al. (2004), as well as Harhoff et. al. (2009) highlight the role of distance with an indication of a potentially negative elasticity on their respective patenting measure. Distance may be negatively related to costs of monitoring the patent, managing business contacts or enforcing patent-law, or simply be related to cultural distance. We therefore use a measure of the distance between two countries' capital cities obtained from the CEPII. We log-transform this series and include the variable $\log(\text{DISTANCE}_{ij})$ in the regressions below.

Though the theoretical model abstracts from trade it has been examined as a determinant of patenting activity previously in the literature: Bosworth (1984) studies in- and out-flows of patent applications to and from the UK, and finds that exports from the UK to a foreign country have a positive effect on UK patenting-activity abroad, whereas imports from abroad have no significant effect on patenting-activity originating from foreign countries to the UK. His findings have been confirmed by Yang and Kuo (2008). Autor et. al. (2016) examine the relationship of US innovations in the manufacturing sector to import competition from China, and conclude a negative effect from raising imports on patenting activity.

To take account of the potential role of trade as an alternative determinant of patent-flows, we employ a measure of bilateral trade between a pair of countries kl at time t that is constructed from data on exports and imports as provided by the DOTS data published by the IMF. Bilateral trade is therefore measured as the sum of half-importer and half-exporter flows. We use a log-transformed version $\log(\text{TRADE}_{kl})$ as an explanatory variable in the regressions below.

Moreover, in order to construct a robustness exercise for the empirical findings derived in this paper by allowing for potential endogeneity of trade we make use of a set of additional variables that provide exclusion restrictions with respect to the potentially endogenous regressor. Therefore, we consider variables standardly employed in the empirical gravity literature (cf. Santos Silva and Tenreyro, 2006): A binary variable indicating colonial ties (COLONY_{kl}, a contiguity-dummy (CONTIG_{kl}), a dummy equal to one if the countries ever were the same country (SMCTRY_{kl}) – all taken from the CEPII, indicator variables for whether two countries share a regional trade agreement (RTA_{klt}) and a common currency (COMCUR_{klt}) both from de Sousa (2012). Moreover we add measures for two countries' bilateral size $\log(\text{SIZE}_{klt}) = \log(\text{GDP}_{kt} + \text{GDP}_{lt})$ and differences in relative factor endowments $\text{RELFAC}_{klt} = \text{abs}[\log(\text{GDPpc}_{kt}) - \log(\text{GDPpc}_{lt})]$ where GDPpc stands for real GDP per capita.

4.3.6 Further Control Variables

Finally, we add a set of further control variables: Measures of net-FDI in- and outflows as a share of GDP at time t at the level of the destination- and origin-country (FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, and FDI-Outlows$_{lt}$ below; cf. MacGarvie, 2005), controls for the investment share in both countries at time t (INVSHARE$_{kt}$ and INVSHARE$_{lt}$ below) The raw-data stem from the WDI[29], and measures of rational forward-looking expectations about GDP in origin and destination log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$ below). Finally, we construct a measure of dissimilarity in the level of globalization for origin and destination at time t from raw-data provided by the KOF on the globalization index suggested in Dreher (2006). The original index consists of three sub-indices for a country's economic, social, and political globalization. We construct the employed dissimilarity-measure as the Euclidean distance between two countries in a given year in terms of these three coordinates. We log-transform the obtained distance and refer to this variable by log(DISSIMI$_{klt}$) below.

4.3.7 Estimation Samples

According to data availability on the distinct measures of inequality, and the other explanatory variables, and as based on the DOCDB-patent-family-specific definitions of the dependent variable, we define the estimation samples as summarized in Table 3. We will refer to the samples where we observe GINInet$_{lt}$ and GINImar$_{lt}$ as the SWIID-samples, and to the samples where we observe top-income-shares (TOP10%-share$_{lt}$, TOP5%-share$_{lt}$, and TOP1%-share$_{lt}$) we refer to as the WID-samples.

We define two different sets of explanatory variables contained in z_{iklt}: A baseline set of regressors $z_{iklt} = z^0_{iklt}$ and an extended set of regressors $z_{iklt} = (z^0_{iklt}, z^1_{iklt})$. Therefore, the baseline set of regressors contains the following variables,

$$z^0_{iklt} = [\ \log(\text{GDP}_{kt}), \log(\text{GDP}_{lt}), \log(\text{POP}_{kt}), \log(\text{POP}_{lt}), \text{GP-INDEX}_{lt}, \\ \text{PATVAL}_{it}, \text{TECHSIM}_{klt}, \text{COMLANG}_{lt}, \log(\text{DIST}_{lt}), \log(\text{TRADE}_{klt})]$$

and besides the variables contained in z^0_{iklt}, the extended set of regressors in addition

[29] These refer to the WDI variables BX.KLT.DINV.WD.GD.ZS, BM.KLT.DINV.WD.GD.ZS and NE.GDI.TOTL.ZS.

includes the following variables

$$z^1_{iklt} = [\log(\text{DISSIMI}_{klt}), \log(\text{NOINVENTORS}_i), \log(\text{ARTpc}_{lt}), \text{FDI-Inflows}_{kt}, \\ \text{FDI-Inflows}_{lt}, \text{FDI-Outflows}_{kt}, \text{FDI-Outlows}_{lt}, \text{INVSHARE}_{kt}, \text{INVSHARE}_{lt}, \\ \log(\text{GDP}_{k(t+1)}), \log(\text{GDP}_{l(t+1)}), \log(\text{GDP}_{k(t+2)}), \log(\text{GDP}_{l(t+2)})].$$

As referred to in Appendix B we have applied missing value imputation for the variables extracted from the WDI assuming missingness at random.

According to these two specification we observe samples ranging from 1,842,972 to 2,874,437 observations, where the SWIID-samples have richer country-coverage as compared to the WID-samples. This information is summarized in Table 3. The SWIID-samples cover 49-53 distinct origin-countries, and 49-50 distinct destinations. The WID-samples cover 47-50 different origins of patent flows and just 17-20 distinct destinations. A full list of country-coverage by sample, origin and destination is given in Appendix B. For the SWIID-samples these numbers refer to 60 (59) distinct countries for the baseline (extended) specification on explanatory variables. For the WID-samples and the baseline (extended) set of control variables there are 51 (48) distinct countries for inequality measure TOP1%-share$_{lt}$, 51 (49) for TOP5%-share$_{lt}$ and TOP10%-share$_{lt}$, respectively.

For all samples considered, we observe patent-flows and the explanatory variables over 23 out of 24 NACE2 Divisions and for each year in 1980-2013.

[Table 3]

4.4 Estimation Results

To assess the empirical implications of the main theoretical results as stated in Corollary 1, by testing the sign restrictions on the coefficients β_2 and β_3 in the relationship stated in equation (13), Tables 4.1-4.5 and Tables 5.1-5.4 contain the estimation results for the weighted count Q^w_{iklt} when controlling for the baseline and the extended set of explanatory variables, respectively. Tables 6.1-6.5 contain a more condensed presentation when using the unweighted count of patent-flows (Q^u_{iklt} as defined above) instead.[30] In these tables, we provide two sorts of inference: Once we cluster standard errors at the level of the innovation i, and once we cluster at the level of the NACE2 Division.

For both SWIID-samples we obtain a statistically highly robust result in favor of what is predicted by Corollary 1: We would not reject that $\beta_2 > 0$ and that $\beta_3 < 0$ for 40 out of 40 different regressions (cf. Tables 4.1-4.2, 5.1-5.2, 6.1-6.2) at a significance level

[30] Robustness checks when considering a different family-defintion (single priority and INPADOC), that will be made available in the Online Appendix to this paper, suggest a qualitatively similar pattern of results as is suggested for the results presented here.

of 0.01 (for either type of inference). For the WID-samples, we find a consistent pattern in favor of the theoretical prediction of Corollary 1 when looking at the top-1% income shares as measure of inequality (cf. Tables 4.5, 5.5, and 6.5, respectively). Significance of the interaction of TOP1%-share$_{lt}$ $*$ log (GDP$_{lt}$) as captured by the coefficient β_3 is not alway given when clustering at the level of the NACE2 Division. For top-10%- and top-5% income shares (TOP10%-share$_{lt}$ and TOP5%-share$_{lt}$; cf. Tables 4.3-4.4, 5.3-5.4, and 6.3-6.4) the results are a bit ambiguous regarding the significance/sign of β_3. The pattern $\beta_1 < 0$ and $\beta_2 > 0$ from above is also strongly supported here.[31] Moreover, we find a highly robust pattern of a negative main-effect of inequality, i.e. $\beta_1 < 0$, and the order of magnitude of the estimated coefficients on the main effect is substantive.

Looking at, e.g., Table 4.1, as we would expect for a correctly specified conditional mean function, the sign and the orders of magnitude of the estimated coefficients for the linear index hardly vary over different models (log-linear, Poisson, negative Binomial, and Gamma). The same pattern generally holds true for all the presented results.

Performing model comparison, the results suggest that for Tables 4.1-4.5 and 5.1-5.5 the log-linear model has the lowest CV error for predicting the underlying dependent variable for a new data-point, and also yields the highest correlation among true and fitted values of the underlying dependent variable for the extended specification (Tables 5.1-5.5). For the baseline specification (Tables 4.1-4.5) the correlation among fitted and true values for the estimated conditional mean is higher for the non-linear models where with respect to both criteria the Poisson model performs best, followed by the negative binomial, and last the Gamma model. Nevertheless, the three non-linear regressions are very close in the quality of fit. Comparing among the baseline (Tables 4.1-4.5) and the extended (Tables 5.1-5.5) specification of explanatory variables, the fit is improved – especially for the linear model – when adding the further regressors. Looking at the other definition of the dependent variable (Tables 6.1-6.5) the log-linear model always has the best fit. The correlation among the fitted and true values of Q^u_{iklt} is higher as compared to using Q^w_{iklt}, but the CV error is smaller for the weighted count. As we will extrapolate from the sample of destination-countries used for estimation in an exercise presented below, good out-of-sample performance is desired most.

Let us briefly consider what the suggested large negative main effect of inequality means quantitatively. Suppose that -log(POP$_{lt}$)β_2 = log(GDP$_{lt}$)β_3, then looking at he log-linear models if one would decrease inequality by 0.01 this would imply that patent-flows would on average increase by 8.6% for the Gini coefficient post-tax (Table 4.1), 7.3% for the Gini coefficient previous to taxes (Table 4.2), 12.5% for the top-10% income

[31] As stated above in Corollary 1 in case of a constant utility from consuming an additional innovative good (i.e. $\partial f(C)/\partial C = 1 \; \forall C$) the interaction of inequality and income vanishes.

share (Table 4.3), for the top-5% income share by 14.5% (Table 4.4), and 16.3% for the top-1% income share (Table 4.5), as suggested by weighted count dependent variable and the baseline set of regressors, and according increases of 11.8% (Table 6.1, left column), 7.2% (Table 6.2, left column), 11.4% (Table 6.3, left column), 13.9% (Table 6.4, left column), and 17.7% (Table 6.5, left column), respectively, for the unweighted count as the dependent variable. For the extended set of control variables and the weighted count we estimate according increases in patent flows due to a decrease in inequality by 0.01 of 11.8% (Table 5.1), 8.6% (Table 5.2), 10.4% (Table 5.2), 11.8% (Table 5.4), and 16.9% (Table 5.5), respectively, and of 16.2% (Table 6.1, right column), 9.1% (Table 6.2, right column), 8.4% (Table 6.3, right column), 9.2% (Table 6.4, right column), and 12.4% (Table 6.5, right column), respectively, for the unweighted count.

[Tables 4.1-4.5, Tables 5.1-5.5, Tables 6.1-6.5]

4.4.1 Inequality and Patent Protection

According to Proposition 2 we would expect that inequality and the level of patent protection in the destination-country interact positively, i.e. a negative effect of inequality on innovation is decreasing with the level of patent protection increasing. Using the same set of specifications as above, we add an interaction term of patent protection (GP-INDEX$_{lt}$) and the employed measure of inequality. We present the according estimation results for the SWIID- and the WID-samples in Tables 7.1-7.5 utilizing the weighted count as the dependent variable. The results suggest a robust pattern in favor of the theoretically predicted relationship. Like above, the results for the unweighted count are qualitatively similar, and will be made available in the Online Appendix. Again we would pick the log-linear and the Poisson models as the best models.

[Tables 7.1-7.5]

4.4.2 The Effect of Inequality on Patent Flows

For a quantitative assessment of the estimation results presented above, we construct a sample of 184 countries where we observe GDP and population for the year 2015. For 115 countries thereof, are also covered by the index of patent protection as provided by Ginarte and Park (2008). Analogously to above, we use non-parametric time-series regression to predict the value of the index for 2015. A complete list of all countries included in the prediction exercises presented here is given in Appendix B1.

In Figure 5 we plot the decision boundary of a zero marginal effect of inequality ($d(I_{lt})$ as defined above in equation (14)) for the estimated log-linear models and the extended set of regressors (Tables 6.1-6.5, second columns) along with other contour lines. The dependent variable is therefore measured on logarithmic scale. The blue and red dots indicate whether a decrease or an increase in inequality would stimulate patent flows and the according country-names are abbreviated by their ISO2 codes.

Translating the displayed pattern of a positive versus a negative effect of inequality in these countries on patent flows, we calculate the probability of a positive effect of an increase in inequality over the underlying 184 countries. For the estimates derived from the SWIID-samples we obtain probabilities of approx. 0.58 and 0.42 for inequality in terms of the Gini coefficient pre- and post-tax, respectively. For the estimates derived from the WID-samples, we calculate probabilities of approx. 0.17, 0.14, and 0.03 for inequality measured by top- 10%-, top-5%-, and top-1%-shares, respectively.

In Figure 6 we plot the decision boundary of a zero marginal effect as function of GDP, population and the level of patent protection as based on the estimation results for the log-linear model and the extended set of control variables (Tables 7.1-7.5, second columns) for a sample of 115 countries and data for the year 2015. From this sample, and for the estimates derived from the SWIID-samples we obtain probabilities of a positive effect of an increase in inequality on patent-flows from approx. 0.73 and 0.49 for inequality in terms of the Gini coefficient pre- and post-tax, respectively. For the estimates derived from the WID-samples, we find according probabilities of approx. 0.06, 0.05, and 0.03 for inequality measured by top- 10%-, top-5%-, and top-1%-shares, respectively.

What pattern is highlighted by the previous two prediction exercises is that inequality post-taxes is too low for a majority of countries to maximize patent flows, and that inequality in terms of top-income-shares is too high for a majority of countries to maximize patent flows. Whereas a high level of patent protection being guaranteed by a country's institutions nets out some of the negative effect of inequality in terms of the Gini coefficient, even a high level of patent protection cannot change the negative effect that an increase in top-income inequality has on patent flows (outside of an incremental range).

[Figure 5, Figure 6]

4.5 Robustness Checks

4.5.1 The Role of Trade

In this robustness check we want to assess whether the results presented above are sensitive to the role of trade. The results presented above in Tables 4.1-4.5 and 5.1-5.5 quite robustly indicate a small positive or a zero effect of trade on patent-flows. In a naive check, we simply drop trade from each of the specifications and find virtually no sensitivity on the estimates of the coefficients β_1, β_2, and β_3. A condensed summary of these results is contained in Appendix B4.

Moreover trade could be endogenous. For a sensitivity analysis when $\log(\text{TRADE}_{klt})$ is allowed to be endogenous in the regression presented in Tables 4.1-4.5 and 5.1-5.5 we apply a control function approach inspired by Wooldridge (1997). Therefore, we firstly regress $\log(\text{TRADE}_{klt})$ at the country-pair-time level on the additional instruments from above, i.e. [$\log(\text{GDPpc}_{kt})$, $\log(\text{GDPpc}_{lt})$, RTA_{klt}, COMCUR_{klt}, SMCTRY_{kl}, COLONY_{kl}, CONTIG_{kl}, $\log(\text{SIZE}_{klt})$, RELFAC_{klt}], and then predict the residual from this regression for all observations in the estimation sample, $\hat{u}_{klt}$. This initial regression is based on 23,301 country-pair-time observations with a linear R^2 of 0.4422.

This residual is then regressed on all the explanatory variables included in the main equation (i.e. on x_{klt}, the additional controls z_{iklt}, and the fixed-effects dummies). The estimated residual from the second regression, $\hat{v}_{iklt}$ is then used as a control function that we add to the explanatory variables. In an alternative exercise that can be found in Appendix B4, we estimate the residual for the first regression by an additive model where the relationship of trade and the instruments is an additive function of non-parametrically estimated functions of each of the explanatory variables. The results presented in the main text are insensitive to using the less-restrictive model.

For inference in the light of the generated regressor $\hat{v}_{iklt}$, we re-estimate $\hat{u}_{iklt}$, $\hat{v}_{iklt}$, and the coefficients of interest over 50 replications from a a clustered bootstrap where for each replication a sample with replacement is drawn form clusters defined by broad sectors j (thus any combination of $iklt$ is retained for a sampled cluster j). For a significance level α, we then compute a two-sided bootstrap-confidence interval as e.g. for $\hat{\beta}_1$ by $[2\hat{\beta}_1 - q^*_{1-\alpha/2}, 2\hat{\beta}_1 - q^*_{\alpha/2}]$, where q^* is the respective quantile of the bootstrap distribution. As suggested by the results in Tables 8.1-8.5 the conclusions derived here are not sensitive to the role of trade, and we would not change the conclusion from above that patent-flows are likely to be increasing or unaffected by the level of bilateral trade among origin- and destination, where at least a negative effect seems very unlikely for our study. However, there are unobservables correlated with trade that have a small negative effect on patent-flows.

[Tables 8.1-8.5]

4.5.2 The Role of Unobserved Heterogeneity

4.5.3 Further Robustness Checks

[TO BE ADDED]

5 Conclusion

This paper has analyzed how inequality can affect innovation through the channel of demand. The two main theoretical results that come out of this analysis are that inequality is more likely beneficial for innovation the larger the size of the population is for a given income and the larger (in terms of cost savings) the innovation step sizes are. We found empirical evidence in line with these model predictions.

Gordon (2016) has argued that there is currently a period of secular stagnation because (unlike in the past) there is a lack of big breakthrough innovations. At the same time, inequality has been increasing in many developing countries. As our analysis suggests that inequality is more likely harmful for innovation the smaller the innovation step sizes are, it therefore provides a possible explanation for the recent slowdown in productivity growth. By increasing the demand for new innovative goods, a more equal distribution of income might therefore, ceteris paribus, lead to more innovation.

Tables and Figures

Figure 1

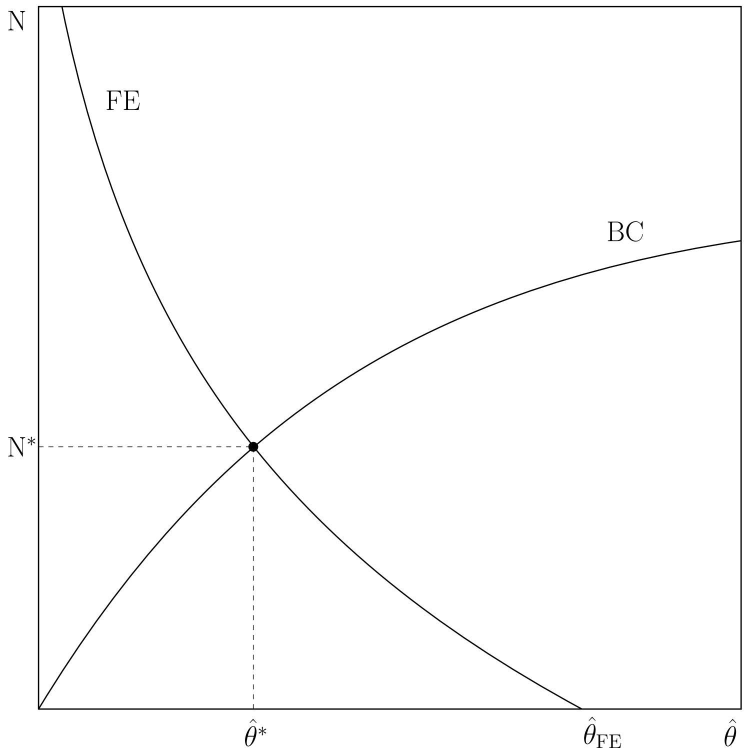
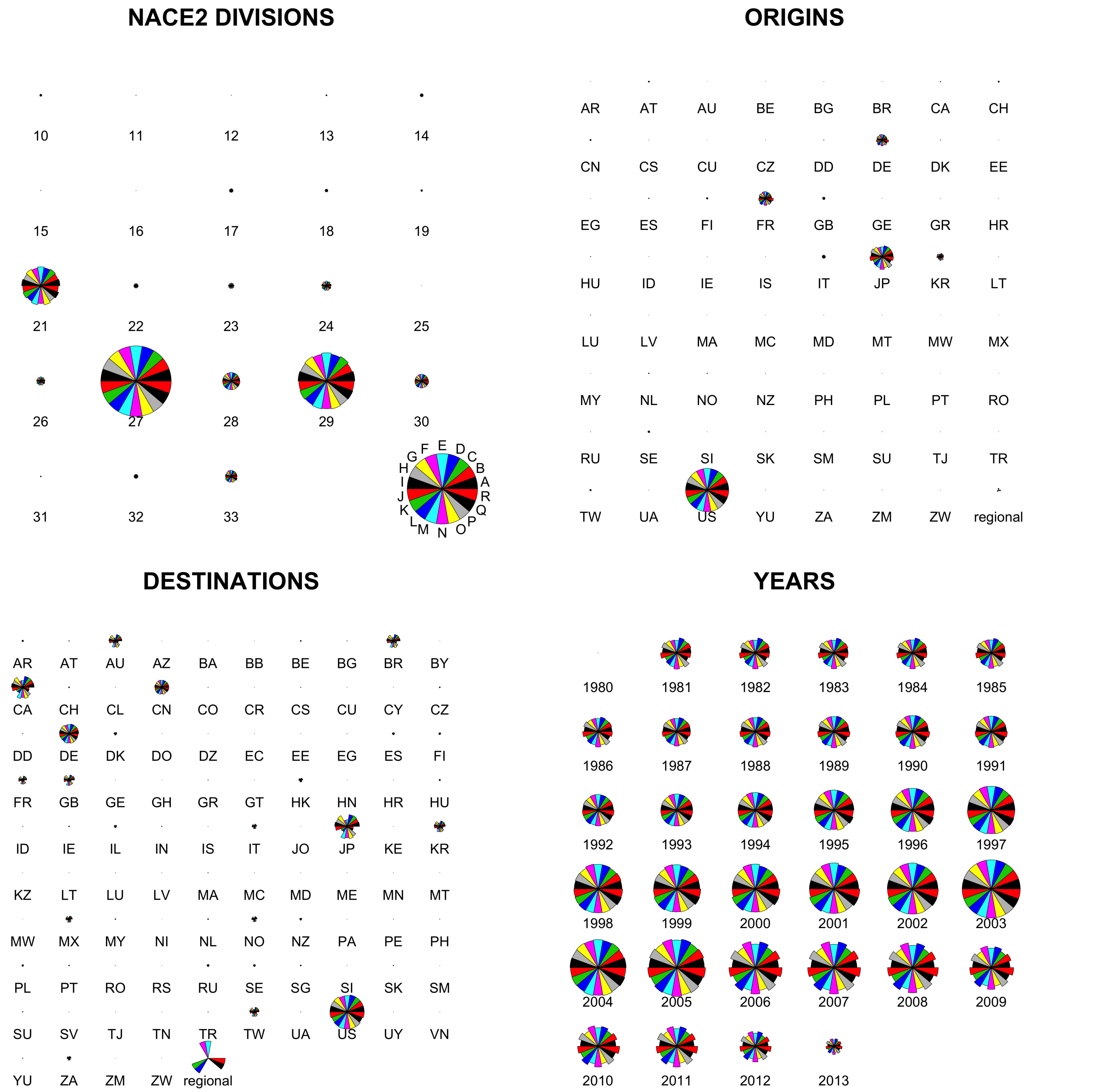


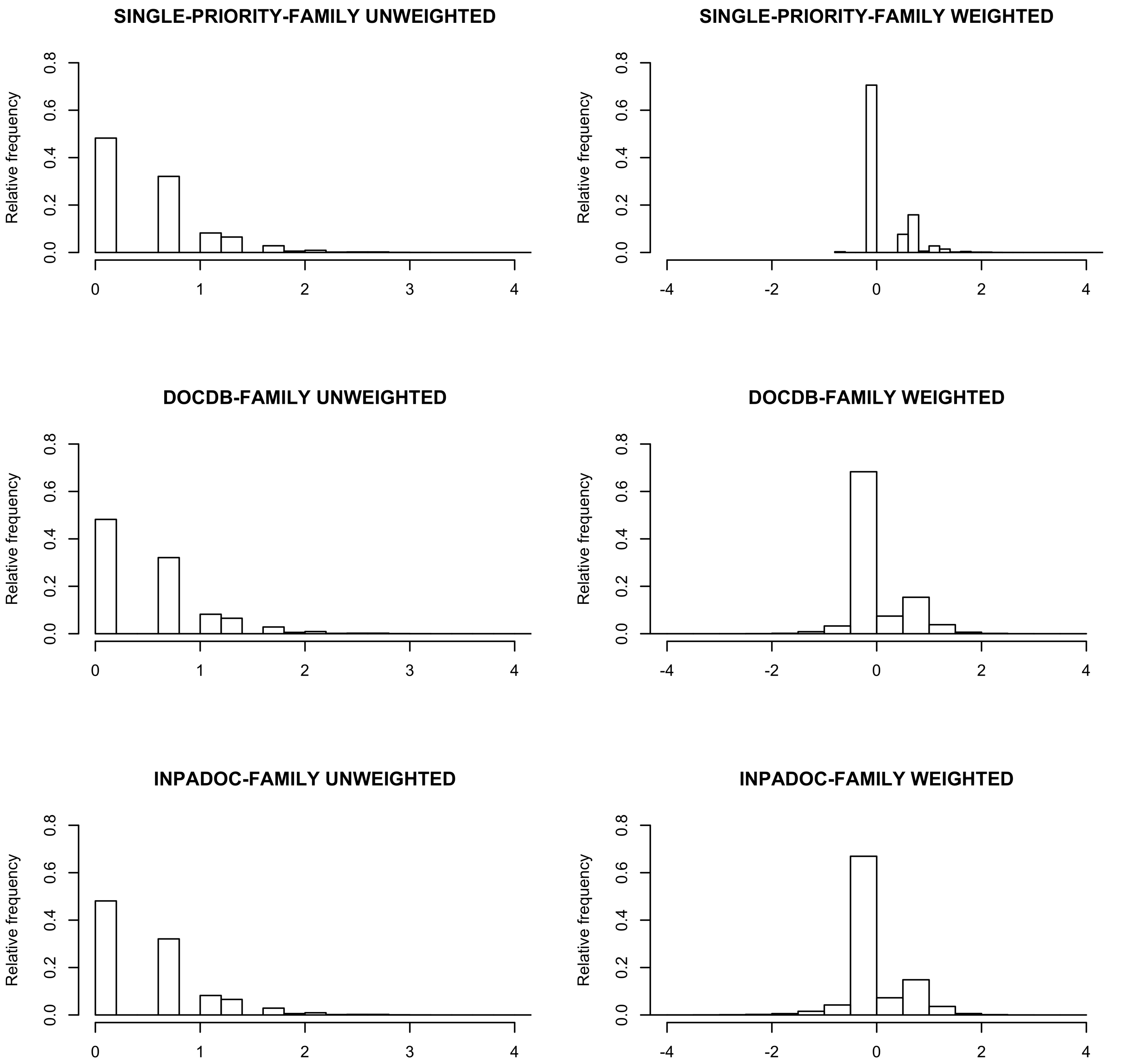
Figure 2. Different definitions of the dependent variable and patterns of patent-flows over NACE2 Manufacturing Divisions, countries of origin and destination, and time



Legend: A. . . unweigted count over PCT- & Paris-Convention subsequent filings without regional applications single-priority-family, **B**. . . weighted count over PCT- & Paris-Convention subsequent filings without regional applications single-priority-family, **C**. . . unw. count over Paris-Convention subsequent filings without regional applications single-priority-fam., **D**. . . w. count over Paris-Convention subsequent filings without regional applications single-priority-fam., **E**. . . unw. count over PCT- & Paris-Convention subsequent filings including regional applications single-priority-fam., **F**. . . w. count over PCT- & Paris-Convention subsequent filings including regional applications single-priority-fam., **G**. . . unw. count over PCT- & Paris-Convention subsequent filings without regional applications DOCDB-fam., *(legend continued on next page)*

Legend continued: **H**. . . w. count over PCT- & Paris-Convention subsequent filings without regional applications DOCDB-fam., **I**. . . unw. count over Paris-Convention subsequent filings without regional applications DOCDB-fam., **J**. . . w. count over Paris-Convention subsequent filings without regional applications DOCDB-fam., **K**. . . unw. count over PCT- & Paris-Convention subsequent filings including regional applications DOCDB-fam., **L**. . . w. count over PCT- & Paris-Convention subsequent filings including regional applications DOCDB-fam., **M**. . . unw. count over PCT- & Paris-Convention subsequent filings without regional applications INPADOC-fam., **N**. . . w. count over PCT- & Paris-Convention subsequent filings without regional applications INPADOC-fam., **O**. . . unw. count over Paris-Convention subsequent filings without regional applications INPADOC-fam., **P**. . . w. count over Paris-Convention subsequent filings without regional applications INPADOC-fam., **Q**. . . unw. count over PCT- & Paris-Convention subsequent filings including regional applications INPADOC-fam., **R**. . . w. count over PCT- & Paris-Convention subsequent filings including regional applications INPADOC-fam.

FIGURE 3. Frequency distributions of patent-flows based on PCT- & Paris-Convention subsequent filings from national applications



Legend: **SINGLE-PRIORITY-FAMILY UNWEIGHTED** (**A** above), **SINGLE-PRIORITY-FAMILY WEIGHTED** (**B** above), **DOCDB-FAMILY UNWEIGHTED** (**G** above), **DOCDB-FAMILY WEIGHTED** (**H** above), **INPADOC-FAMILY UNWEIGHTED** (**M** above), **INPADOC-FAMILY WEIGHTED** (**N** above)

TABLE 1. Descriptive statistics for data on dependent variable as used for main-results (patent-flows based on PCT- & Paris-Convention subsequent filings from national applications with different family definitions and weighting schemes)

	Single-priority-family		DOCDB-family		INPADOC-family	
Obs.	3,735,424		3,731,384		3,710,940	
	UNWEIGHTED	WEIGHTED	UNWEIGHTED	WEIGHTED	UNWEIGHTED	WEIGHTED
Mean	1.998	1.346	2.000	1.295	2.011	1.263
Variance	2.727	0.583	2.765	0.548	2.916	0.539
Min.	1	0.5	1	0.007	1	0.001
Max.	160	80	160	52.50	160	52.50

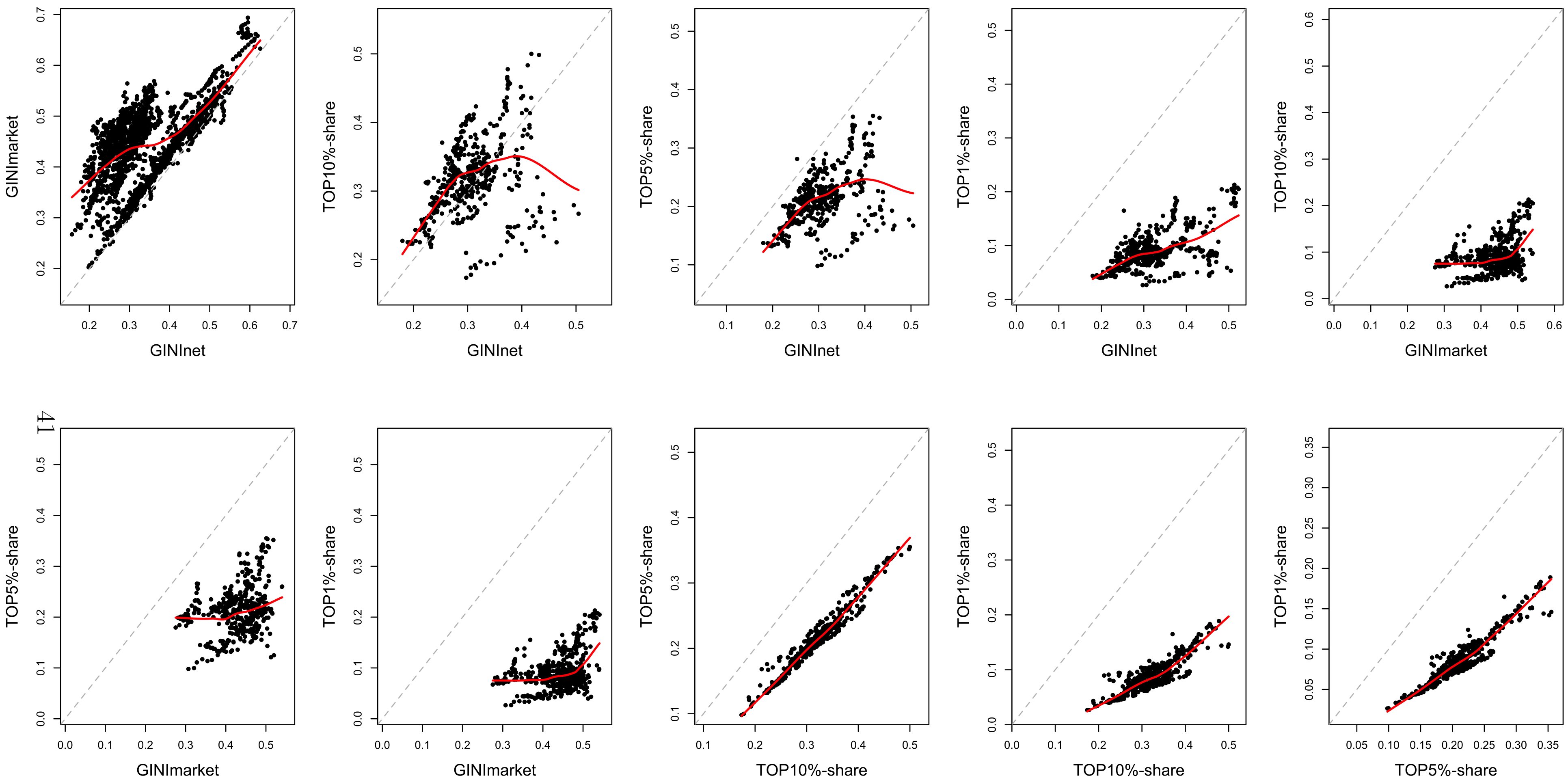
Note: Different family definitions as levels of aggregation for measuring distinct innovations as described in the text. There are 1,503,641 distinct innovations for the single-priority-family, 1,501,514 innovations for the DOCDB-family, and 1,481,657 for the INPADOC-family, respectively.

TABLE 2. Analysis of variance for log-transformed measures of the dependent variable as used for main-results (patent-flows based on PCT- & Paris-Convention subsequent filings from national applications with different family definitions and weighting schemes)

		Single-priority-family		DOCDB-family		INPADOC-family	
	DF	UNWEIGHTED	WEIGHTED	UNWEIGHTED	WEIGHTED	UNWEIGHTED	WEIGHTED
NACE2-Division δ_o	22	27532.5***	22823.3***	27132.3***	22823.3***	25606.5***	21159.9***
Origin-country δ_k	61	297.6***	350.4***	299.7***	350.4***	310.8***	523.6***
Destin.-country δ_l	92	1991.8***	4328.8***	1963.1***	4328.8***	1852.4***	4116.2***
Year δ_t	33	995.4***	306.7***	982.8***	306.7***	943.6***	346.9***
Total sum squared		10186004.8	2177909.1	$> 10^{14}$	2045997.6	10821226.1	2000099.8
Residual sum squared		8228598.3	1695251.2	8355166.5	1615692.8	8843173.2	1592513.4
R^2		0.1922	0.1902	0.1902	0.2103	0.1828	0.2038

lote: ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively (errors assumed to be iid over index $iklt$).

Figure 4. Marginal relationships among inequality measures for as observed for any destination country over time (1980-2013)



Note: Smooth red curves represent a locally weighted scatterplot smoother (LOESS-regression) of the inequality measure on the ordinate on the inequality measure on the abscissa. For orientation, a 45-degree line was added to each plot (dashed gray).

Table 3. Estimation samples with respect to different inequality measures

	SWIID-samples	WID-samples		
	Gini pre- & post-tax	top-10%-share	top-5%-share	top-1%-share
BASELINE				
Observations	2,874,437	1,919,377	1,913,219	1,931,973
Innovations	1,336,544	1,189,422	1,189,189	1,191,443
NACE2 Divisons	23	23	23	23
Origins	53	50	50	50
Destinations	50	17	17	20
Periods of time	1980-2013	1980-2013	1980-2013	1980-2013
EXTENDED				
Observations	2,737,773	1,842,972	1,837,315	1,855,458
Innovations	1,278,296	1,138,413	1,138,206	1,140,416
NACE2 Divisons	23	23	23	23
Origins	49	47	47	47
Destinations	49	17	17	20
Periods of time	1980-2013	1980-2013	1980-2013	1980-2013

Table 4.1. Results baseline specification for SWIID Gini pre-tax (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$	Poisson$^{(1)}/^{(2)}$	Neg. bin.$^{(1)}/^{(2)}$	Gamma$^{(1)}/^{(2)}$
GINInet$_{lt}$	$-8.206^{***/***}$	$-8.252^{***/***}$	$-9.340^{***/***}$	$-9.815^{***/***}$
	(0.214)	(0.265)	(0.237)	(0.227)
log(POP$_{lt}$)*GINInet$_{lt}$	$0.688^{***/***}$	$0.731^{***/***}$	$0.793^{***/***}$	$0.810^{***/***}$
	(0.014)	(0.017)	(0.015)	(0.015)
log(GDP$_{lt}$)*GINInet$_{lt}$	$-0.100^{***/***}$	$-0.122^{***/***}$	$-0.124^{***/***}$	$-0.119^{***/***}$
	(0.003)	(0.004)	(0.003)	(0.003)
log(GDP$_{kt}$)	$0.044^{***/***}$	$0.042^{***/***}$	$0.041^{***/***}$	$0.040^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.003)
log(GDP$_{lt}$)	$0.051^{***/***}$	$0.060^{***/***}$	$0.061^{***/***}$	$0.058^{***/***}$
	(0.001)	(0.002)	(0.002)	(0.002)
log(POP$_{kt}$)	$-0.618^{***/***}$	$-0.280^{***/**}$	$-0.254^{***/**}$	$-0.245^{***/**}$
	(0.014)	(0.018)	(0.016)	(0.015)
log(POP$_{lt}$)	-0.010	$0.087^{***/}$	$0.048^{***/}$	$0.029^{**/}$
	(0.009)	(0.011)	(0.010)	(0.010)
GP-INDEX$_{lt}$	$0.039^{***/***}$	$0.045^{***/***}$	$0.046^{***/***}$	$0.045^{***/***}$
	(0.001)	(0.002)	(0.001)	(0.001)
PATVAL$_{it}$	$0.057^{***/***}$	$0.082^{***/***}$	$0.076^{***/***}$	$0.073^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-0.429^{***/**}$	$-0.516^{***/***}$	$-0.383^{***/***}$	$-0.306^{***/**}$
	(0.042)	(0.054)	(0.048)	(0.046)
COMLANG$_{lt}$	$-0.015^{***/*}$	-0.002	$-0.006^{***/}$	$-0.010^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(DIST$_{lt}$)	$0.011^{***/*}$	$0.011^{***/}$	$0.010^{***/}$	$0.010^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.020^{***/***}$	$0.023^{***/***}$	$0.023^{***/***}$	$0.023^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
χ^2 NACE2 Division	61005.8 $^{***/***}$	57569.8 $^{***/***}$	62510.3 $^{***/***}$	63751.9 $^{***/***}$
χ^2 Origin	16934.3 $^{***/***}$	9975.1 $^{***/***}$	11266.4 $^{***/***}$	11775.9 $^{***/***}$
χ^2 Destination	255791.2 $^{***/***}$	206635.4 $^{***/***}$	241413.9 $^{***/***}$	254262.6 $^{***/***}$
χ^2 Time	5283.3 $^{***/***}$	4241.2 $^{***/***}$	5355.6 $^{***/***}$	5868.6 $^{***/***}$
Corr(true,fitted)	0.5091	0.5486	0.5465	0.5450
CV error	0.1367	0.4115	0.4130	0.4142

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 4.2. Results baseline specification for SWIID Gini post-tax (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$	Poisson$^{(1)}/^{(2)}$	Neg. bin.$^{(1)}/^{(2)}$	Gamma$^{(1)}/^{(2)}$
GINImar_{lt}	$-7.068^{***/***}$	$-6.740^{***/***}$	$-7.012^{***/***}$	$-7.041^{***/***}$
	(0.147)	(0.181)	(0.161)	(0.154)
$\log(\text{POP}_{lt})$*GINImar_{lt}	$0.653^{***/***}$	$0.680^{***/***}$	$0.673^{***/***}$	$0.650^{***/***}$
	(0.010)	(0.013)	(0.011)	(0.011)
$\log(\text{GDP}_{lt})$*GINImar_{lt}	$-0.138^{***/***}$	$-0.163^{***/***}$	$-0.152^{***/***}$	$-0.139^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.003)
$\log(\text{GDP}_{kt})$	$0.045^{***/***}$	$0.043^{***/***}$	$0.041^{***/***}$	$0.041^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.003)
$\log(\text{GDP}_{lt})$	$0.077^{***/***}$	$0.090^{***/***}$	$0.083^{***/***}$	$0.076^{***/***}$
	(0.001)	(0.002)	(0.002)	(0.001)
$\log(\text{POP}_{kt})$	$-0.629^{***/***}$	$-0.294^{***/**}$	$-0.271^{***/**}$	$-0.263^{***/**}$
	(0.014)	(0.018)	(0.016)	(0.015)
$\log(\text{POP}_{lt})$	$0.073^{***/}$	$0.176^{***/***}$	$0.163^{***/**}$	$0.154^{***/**}$
	(0.009)	(0.011)	(0.010)	(0.009)
GP-INDEX_{lt}	$0.023^{***/***}$	$0.028^{***/***}$	$0.026^{***/***}$	$0.025^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
PATVAL_{it}	$0.057^{***/***}$	$0.082^{***/***}$	$0.076^{***/***}$	$0.073^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM_{klt}	$0.102^{*/}$	0.062	$0.134^{**/}$	$0.160^{***/}$
	(0.042)	(0.055)	(0.049)	(0.046)
COMLANG_{lt}	$-0.015^{***/*}$	-0.002	$-0.006^{***/}$	$-0.010^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{DIST}_{lt})$	$0.011^{***/**}$	$0.011^{***/}$	$0.011^{***/*}$	$0.010^{***/*}$
	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{TRADE}_{klt})$	$0.021^{***/***}$	$0.024^{***/***}$	$0.024^{***/***}$	$0.024^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
χ^2 NACE2 Division	61012.5 $^{***/***}$	57620.5 $^{***/***}$	62513.3 $^{***/***}$	63750.3 $^{***/***}$
χ^2 Origin	16773.4 $^{***/***}$	9907.9 $^{***/***}$	11115.4 $^{***/***}$	11603.0 $^{***/***}$
χ^2 Destination	239709.3 $^{***/***}$	198516.0 $^{***/***}$	234879.9 $^{***/***}$	249108.2 $^{***/***}$
χ^2 Time	5091.8 $^{***/***}$	4464.3 $^{***/***}$	5327.4 $^{***/***}$	5597.1 $^{***/***}$
Corr(true,fitted)	0.5090	0.5487	0.5467	0.5452
CV error	0.1367	0.4114	0.4130	0.4141

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 4.3. Results baseline specification for WID top-10%-share (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$	Poisson$^{(1)}/^{(2)}$	Neg. bin.$^{(1)}/^{(2)}$	Gamma$^{(1)}/^{(2)}$
TOP10%-share$_{lt}$	$-11.771^{***/***}$	$-14.057^{***/***}$	$-12.707^{***/***}$	$-11.694^{***/***}$
	(0.230)	(0.301)	(0.273)	(0.259)
log(POP$_{lt}$)*TOP10%-share$_{lt}$	$0.687^{***/***}$	$0.799^{***/***}$	$0.759^{***/***}$	$0.725^{***/***}$
	(0.017)	(0.022)	(0.020)	(0.019)
log(GDP$_{lt}$)*TOP10%-share$_{lt}$	$-0.027^{***/}$	-0.016	$-0.037^{***/}$	$-0.051^{***/}$
	(0.008)	(0.010)	(0.009)	(0.009)
log(GDP$_{kt}$)	$0.057^{***/***}$	$0.057^{***/***}$	$0.055^{***/***}$	$0.054^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.002)
log(GDP$_{lt}$)	$0.033^{***/**}$	$0.038^{***/**}$	$0.031^{***/*}$	$0.025^{***/}$
	(0.003)	(0.004)	(0.004)	(0.004)
log(POP$_{kt}$)	$-0.907^{***/***}$	$-0.609^{***/***}$	$-0.597^{***/***}$	$-0.593^{***/***}$
	(0.015)	(0.020)	(0.017)	(0.016)
log(POP$_{lt}$)	$-0.055^{***/}$	$0.040^{*/}$	$0.050^{**/}$	$0.058^{**/}$
	(0.016)	(0.020)	(0.018)	(0.018)
GP-INDEX$_{lt}$	0.004	$0.024^{***/}$	$0.018^{***/}$	$0.013^{***/}$
	(0.002)	(0.003)	(0.002)	(0.002)
PATVAL$_{it}$	$0.054^{***/***}$	$0.077^{***/***}$	$0.072^{***/***}$	$0.069^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-1.268^{***/***}$	$-1.562^{***/***}$	$-1.361^{***/***}$	$-1.223^{***/***}$
	(0.051)	(0.072)	(0.065)	(0.061)
COMLANG$_{lt}$	$-0.027^{***/***}$	$-0.006^{**/}$	$-0.011^{***/*}$	$-0.015^{***/**}$
	(0.001)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.009^{***/***}$	$-0.012^{***/***}$	$-0.013^{***/***}$	$-0.013^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.005^{***/*}$	$0.006^{***/}$	$0.005^{***/}$	$0.005^{***/*}$
	(0.001)	(0.001)	(0.001)	(0.001)
χ^2 NACE2 Division	53188.8 $^{***/***}$	46340.4 $^{***/***}$	49912.3 $^{***/***}$	50625.9 $^{***/***}$
χ^2 Origin	15966.4 $^{***/***}$	8972.9 $^{***/***}$	10335.1 $^{***/***}$	10904.0 $^{***/***}$
χ^2 Destination	108584.3 $^{***/***}$	76463.4 $^{***/***}$	91555.9 $^{***/***}$	97493.5 $^{***/***}$
χ^2 Time	15995.2 $^{***/***}$	12800.6 $^{***/***}$	14089.5 $^{***/***}$	14480.1 $^{***/***}$
Corr(true,fitted)	0.5017	0.5332	0.5315	0.5301
CV error	0.1156	0.3309	0.3320	0.3328

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 4.4. Results baseline specification for WID top-5%-share (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$	Poisson$^{(1)}/^{(2)}$	Neg. bin.$^{(1)}/^{(2)}$	Gamma$^{(1)}/^{(2)}$
TOP5%-share$_{lt}$	$-13.526^{***/***}$	$-16.376^{***/***}$	$-14.638^{***/***}$	$-13.381^{***/***}$
	(0.262)	(0.341)	(0.309)	(0.293)
log(POP$_{lt}$)*TOP5%-share$_{lt}$	$0.884^{***/***}$	$0.993^{***/***}$	$0.948^{***/***}$	$0.914^{***/***}$
	(0.019)	(0.025)	(0.022)	(0.021)
log(GDP$_{lt}$)*TOP5%-share$_{lt}$	$-0.101^{***/}$	$-0.067^{***/}$	$-0.098^{***/}$	$-0.119^{***/}$
	(0.011)	(0.014)	(0.013)	(0.013)
log(GDP$_{kt}$)	$0.059^{***/***}$	$0.058^{***/***}$	$0.057^{***/***}$	$0.056^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.002)
log(GDP$_{lt}$)	$0.018^{***/}$	$0.023^{***/}$	$0.015^{***/}$	$0.009^{*/}$
	(0.003)	(0.004)	(0.004)	(0.004)
log(POP$_{kt}$)	$-0.894^{***/***}$	$-0.593^{***/***}$	$-0.582^{***/***}$	$-0.579^{***/***}$
	(0.015)	(0.020)	(0.017)	(0.016)
log(POP$_{lt}$)	$0.046^{**/}$	$0.144^{***/**}$	$0.150^{***/**}$	$0.153^{***/**}$
	(0.015)	(0.019)	(0.018)	(0.017)
GP-INDEX$_{lt}$	$0.009^{***/}$	$0.032^{***/*}$	$0.024^{***/*}$	$0.019^{***/}$
	(0.002)	(0.003)	(0.002)	(0.002)
PATVAL$_{it}$	$0.053^{***/***}$	$0.077^{***/***}$	$0.072^{***/***}$	$0.069^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-1.253^{***/***}$	$-1.530^{***/***}$	$-1.330^{***/***}$	$-1.192^{***/***}$
	(0.051)	(0.072)	(0.064)	(0.060)
COMLANG$_{lt}$	$-0.028^{***/***}$	$-0.006^{***/}$	$-0.012^{***/*}$	$-0.016^{***/***}$
	(0.001)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.011^{***/***}$	$-0.013^{***/***}$	$-0.014^{***/***}$	$-0.015^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.005^{***/*}$	$0.004^{***/}$	$0.004^{***/}$	$0.004^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
χ^2 NACE2 Division	52933.9 $^{***/***}$	46052.3 $^{***/***}$	49601.0 $^{***/***}$	50305.2 $^{***/***}$
χ^2 Origin	15811.8 $^{***/***}$	8878.5 $^{***/***}$	10247.5 $^{***/***}$	10833.3 $^{***/***}$
χ^2 Destination	106728.1 $^{***/***}$	75060.1 $^{***/***}$	89873.6 $^{***/***}$	95784.1 $^{***/***}$
χ^2 Time	16087.4 $^{***/***}$	12622.8 $^{***/***}$	13926.9 $^{***/***}$	14348.3 $^{***/***}$
Corr(true,fitted)	0.5022	0.5332	0.5315	0.5301
CV error	0.1153	0.3291	0.3302	0.3310

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 4.5. Results baseline specification for WID top-1%-share (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$	Poisson$^{(1)}/^{(2)}$	Neg. bin.$^{(1)}/^{(2)}$	Gamma$^{(1)}/^{(2)}$
TOP1%-share$_{lt}$	$-15.104^{***/***}$	$-19.828^{***/***}$	$-17.175^{***/***}$	$-15.301^{***/***}$
	(0.413)	(0.532)	(0.481)	(0.458)
log(POP$_{lt}$)*TOP1%-share$_{lt}$	$1.328^{***/***}$	$1.446^{***/***}$	$1.377^{***/***}$	$1.325^{***/***}$
	(0.028)	(0.036)	(0.032)	(0.031)
log(GDP$_{lt}$)*TOP1%-share$_{lt}$	$-0.344^{***/***}$	$-0.251^{***/**}$	$-0.295^{***/**}$	$-0.324^{***/***}$
	(0.022)	(0.028)	(0.025)	(0.024)
log(GDP$_{kt}$)	$0.059^{***/***}$	$0.058^{***/***}$	$0.057^{***/***}$	$0.056^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.002)
log(GDP$_{lt}$)	$0.021^{***/}$	$0.024^{***/*}$	$0.017^{***/}$	$0.013^{***/}$
	(0.003)	(0.004)	(0.004)	(0.004)
log(POP$_{kt}$)	$-0.916^{***/***}$	$-0.610^{***/***}$	$-0.600^{***/***}$	$-0.598^{***/***}$
	(0.015)	(0.020)	(0.017)	(0.016)
log(POP$_{lt}$)	$0.170^{***/***}$	$0.273^{***/***}$	$0.257^{***/***}$	$0.244^{***/***}$
	(0.015)	(0.019)	(0.017)	(0.017)
GP-INDEX$_{lt}$	-0.001	$0.024^{***/*}$	$0.016^{***/}$	$0.010^{***/}$
	(0.002)	(0.003)	(0.002)	(0.002)
PATVAL$_{it}$	$0.054^{***/***}$	$0.077^{***/***}$	$0.072^{***/***}$	$0.069^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-1.191^{***/***}$	$-1.384^{***/***}$	$-1.210^{***/***}$	$-1.088^{***/***}$
	(0.052)	(0.072)	(0.064)	(0.061)
COMLANG$_{lt}$	$-0.026^{***/***}$	$-0.005^{**/}$	$-0.011^{***/*}$	$-0.015^{***/**}$
	(0.001)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.009^{***/***}$	$-0.012^{***/***}$	$-0.013^{***/***}$	$-0.013^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.006^{***/*}$	$0.006^{***/}$	$0.005^{***/*}$	$0.005^{***/*}$
	(0.001)	(0.001)	(0.001)	(0.001)
χ^2 NACE2 Division	53759.8 $^{***/***}$	47035.1 $^{***/***}$	50673.7 $^{***/***}$	51407.7 $^{***/***}$
χ^2 Origin	15889.8 $^{***/***}$	8870.6 $^{***/***}$	10232.2 $^{***/***}$	10879.8 $^{***/***}$
χ^2 Destination	110042.4 $^{***/***}$	78483.6 $^{***/***}$	94010.8 $^{***/***}$	100332.6 $^{***/***}$
χ^2 Time	14659.2 $^{***/***}$	11438.0 $^{***/***}$	12665.2 $^{***/***}$	13069.4 $^{***/***}$
Corr(true,fitted)	0.5025	0.5342	0.5325	0.5312
CV error	0.1159	0.3310	0.3321	0.3329

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

TABLE 5.1. Results extended specification for SWIID Gini pre-tax (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}$/$^{(2)}$	Poisson$^{(1)}$/$^{(2)}$	Neg. bin.$^{(1)}$/$^{(2)}$	Gamma$^{(1)}$/$^{(2)}$
GINInet$_{lt}$	−11.173***/***	−11.421***/***	−12.276***/***	−12.546***/***
	(0.206)	(0.302)	(0.270)	(0.258)
log(POP$_{lt}$)*GINInet$_{lt}$	0.796***/***	0.856***/***	0.908***/***	0.918***/***
	(0.013)	(0.019)	(0.017)	(0.016)
log(GDP$_{lt}$)*GINInet$_{lt}$	−0.064***/***	−0.087***/***	−0.090***/***	−0.088***/***
	(0.003)	(0.004)	(0.004)	(0.003)
log(GDP$_{kt}$)	−0.018*/	−0.023	−0.023*/	−0.024*/
	(0.009)	(0.012)	(0.011)	(0.010)
log(GDP$_{lt}$)	0.030***/**	0.044***/***	0.043***/***	0.041***/***
	(0.002)	(0.003)	(0.003)	(0.003)
log(POP$_{kt}$)	0.285***/***	0.400***/***	0.393***/***	0.387***/***
	(0.012)	(0.018)	(0.016)	(0.015)
log(POP$_{lt}$)	0.011	0.080***/*	0.051***/	0.038***/
	(0.009)	(0.013)	(0.011)	(0.011)
GP-INDEX$_{lt}$	0.029***/***	0.021***/***	0.028***/***	0.030***/***
	(0.001)	(0.002)	(0.001)	(0.001)
PATVAL$_{it}$	0.061***/***	0.086***/***	0.079***/***	0.076***/***
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	−0.749***/***	−0.806***/***	−0.654***/***	−0.581***/***
	(0.038)	(0.056)	(0.050)	(0.047)
COMLANG$_{lt}$	0.005***/	0.016***/*	0.011***/	0.008***/
	(0.001)	(0.001)	(0.001)	(0.001)
log(DIST$_{lt}$)	0.003***/	0.007***/	0.005***/	0.004***/
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.014***/***	0.021***/***	0.019***/***	0.017***/***
	(0.001)	(0.001)	(0.001)	(0.001)
log(DISSIMI$_{klt}$)	0.003***/**	0.006***/***	0.006***/***	0.005***/***
	(0.000)	(0.001)	(0.001)	(0.001)
log(NOINVENTORS$_{i}$)	0.020***/***	0.025***/***	0.025***/***	0.025***/***
	(0.001)	(0.001)	(0.001)	(0.001)
log(ARTpc$_{lt}$)	0.036***/***	0.045***/***	0.043***/***	0.041***/***
	(0.001)	(0.001)	(0.001)	(0.001)
FDI-Inflows$_{kt}$	0.045*/	0.013	0.039	0.052*/
	(0.022)	(0.031)	(0.028)	(0.026)
FDI-Inflows$_{lt}$	0.047**/**	0.037	0.057**/*	0.068**/***
	(0.016)	(0.024)	(0.022)	(0.021)
FDI-Outflows$_{kt}$	−0.043	0.002	−0.033	−0.048
	(0.022)	(0.032)	(0.029)	(0.027)
FDI-Outlows$_{lt}$	0.142***/***	0.211***/***	0.201***/***	0.198***/***
	(0.016)	(0.025)	(0.022)	(0.021)
INVSHARE$_{kt}$	−0.270***/**	−0.347***/**	−0.326***/**	−0.307***/**
	(0.019)	(0.029)	(0.025)	(0.023)
INVSHARE$_{lt}$	0.020	0.000	0.006	0.010
	(0.011)	(0.018)	(0.016)	(0.015)
log(GDP$_{k(t+1)}$)	0.004	0.007	0.005	0.004
	(0.011)	(0.015)	(0.014)	(0.013)
log(GDP$_{l(t+1)}$)	−0.013***/***	−0.013**/***	−0.011**/**	−0.011**/**
	(0.003)	(0.004)	(0.004)	(0.004)
log(GDP$_{k(t+2)}$)	0.049***/**	0.065***/**	0.063***/**	0.061***/**
	(0.009)	(0.013)	(0.012)	(0.012)
log(GDP$_{l(t+2)}$)	0.015***/***	0.009***/***	0.011***/***	0.013***/***
	(0.002)	(0.002)	(0.002)	(0.002)
χ^2 NACE2 Division	96149.9 ***/***	65095.6 ***/***	73485.1 ***/***	75637.3 ***/***
χ^2 Origin	27741.3 ***/***	13534.3 ***/***	16851.6 ***/***	18846.0 ***/***
χ^2 Destination	224587.7 ***/***	141448.2 ***/***	167781.5 ***/***	177864.2 ***/***
χ^2 Time	7960.8 ***/***	5130.5 ***/***	6103.1 ***/***	6547.3 ***/***
Corr(true,fitted)	0.5652	0.5634	0.5610	0.5594
CV error	0.0988	0.4182	0.4202	0.4215

$^{(1)}$/$^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 5.2. Results extended specification for SWIID Gini post-tax (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$	Poisson$^{(1)}/^{(2)}$	Neg. bin.$^{(1)}/^{(2)}$	Gamma$^{(1)}/^{(2)}$
GINImar_{lt}	$-8.288^{***/***}$	$-8.263^{***/***}$	$-8.342^{***/***}$	$-8.263^{***/***}$
	(0.129)	(0.188)	(0.167)	(0.159)
$\log(\text{POP}_{lt})*\text{GINImar}_{lt}$	$0.709^{***/***}$	$0.763^{***/***}$	$0.742^{***/***}$	$0.714^{***/***}$
	(0.009)	(0.014)	(0.012)	(0.011)
$\log(\text{GDP}_{lt})*\text{GINImar}_{lt}$	$-0.126^{***/***}$	$-0.157^{***/***}$	$-0.143^{***/***}$	$-0.131^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.003)
$\log(\text{GDP}_{kt})$	$-0.018^{*/}$	-0.023	$-0.023^{*/}$	$-0.024^{*/*}$
	(0.009)	(0.012)	(0.011)	(0.010)
$\log(\text{GDP}_{lt})$	$0.069^{***/***}$	$0.088^{***/***}$	$0.078^{***/***}$	$0.070^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.002)
$\log(\text{POP}_{kt})$	$0.264^{***/***}$	$0.376^{***/***}$	$0.364^{***/***}$	$0.355^{***/***}$
	(0.012)	(0.018)	(0.016)	(0.015)
$\log(\text{POP}_{lt})$	$0.140^{***/***}$	$0.203^{***/***}$	$0.205^{***/***}$	$0.205^{***/***}$
	(0.008)	(0.012)	(0.011)	(0.010)
GP-INDEX_{lt}	$0.012^{***/***}$	$0.005^{**/}$	$0.009^{***/}$	$0.011^{***/*}$
	(0.001)	(0.002)	(0.001)	(0.001)
PATVAL_{it}	$0.061^{***/***}$	$0.086^{***/***}$	$0.079^{***/***}$	$0.076^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM_{klt}	$-0.189^{***/**}$	$-0.199^{***/*}$	-0.096	-0.066
	(0.039)	(0.057)	(0.050)	(0.047)
COMLANG_{lt}	$0.006^{***/}$	$0.017^{***/**}$	$0.012^{***/}$	$0.009^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{DIST}_{lt})$	$0.004^{***/}$	$0.006^{***/}$	$0.005^{***/}$	$0.004^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{TRADE}_{klt})$	$0.015^{***/***}$	$0.021^{***/***}$	$0.019^{***/***}$	$0.018^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{DISSIMI}_{klt})$	$0.006^{***/***}$	$0.009^{***/***}$	$0.008^{***/***}$	$0.008^{***/***}$
	(0.000)	(0.001)	(0.001)	(0.001)
$\log(\text{NOINVENTORS}_{i})$	$0.020^{***/***}$	$0.025^{***/***}$	$0.025^{***/***}$	$0.025^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{ARTpc}_{lt})$	$0.029^{***/***}$	$0.036^{***/***}$	$0.034^{***/***}$	$0.032^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
FDI-Inflows_{kt}	$0.052^{*/}$	0.020	0.046	$0.060^{*/}$
	(0.022)	(0.031)	(0.028)	(0.026)
FDI-Inflows_{lt}	$0.018^{/}$	0.017	0.028	0.032
	(0.016)	(0.024)	(0.022)	(0.021)
FDI-Outflows_{kt}	$-0.046^{*/}$	0.000	-0.036	-0.052
	(0.022)	(0.032)	(0.029)	(0.027)
FDI-Outlows_{lt}	$0.154^{***/***}$	$0.214^{***/***}$	$0.195^{***/***}$	$0.189^{***/***}$
	(0.016)	(0.025)	(0.022)	(0.021)
INVSHARE_{kt}	$-0.312^{***/**}$	$-0.392^{***/**}$	$-0.371^{***/**}$	$-0.350^{***/**}$
	(0.020)	(0.029)	(0.025)	(0.023)
INVSHARE_{lt}	-0.015	$-0.050^{**/}$	$-0.035^{*/}$	-0.024
	(0.011)	(0.018)	(0.015)	(0.014)
$\log(\text{GDP}_{k(t+1)})$	0.005	0.008	0.006	0.005
	(0.011)	(0.015)	(0.014)	(0.013)
$\log(\text{GDP}_{l(t+1)})$	$-0.009^{**/**}$	-0.004	-0.006	$-0.008^{*/*}$
	(0.003)	(0.004)	(0.004)	(0.004)
$\log(\text{GDP}_{k(t+2)})$	$0.049^{***/**}$	$0.065^{***/**}$	$0.063^{***/**}$	$0.061^{***/**}$
	(0.009)	(0.013)	(0.012)	(0.012)
$\log(\text{GDP}_{l(t+2)})$	$0.011^{***/***}$	$0.001^{/}$	$0.006^{**/**}$	$0.009^{***/***}$
	(0.002)	(0.002)	(0.002)	(0.002)
χ^2 NACE2 Division	96182.5 $^{***/***}$	65169.8 $^{***/***}$	73457.5 $^{***/***}$	75593.9 $^{***/***}$
χ^2 Origin	27293.7 $^{***/***}$	13354.1 $^{***/***}$	16466.8 $^{***/***}$	18327.5 $^{***/***}$
χ^2 Destination	217773.1 $^{***/***}$	136953.4 $^{***/***}$	164750.2 $^{***/***}$	175824.6 $^{***/***}$
χ^2 Time	8232.3 $^{***/***}$	5203.6 $^{***/***}$	6171.8 $^{***/***}$	6564.5 $^{***/***}$
Corr(true,fitted)	0.5653	0.5636	0.5612	0.5596
CV error	0.0987	0.4181	0.4200	0.4214

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 5.3. Results extended specification for WID top-10%-share (dep. var. Q^w_{iklt})

	Log-linear(1)/(2)	Poisson(1)/(2)	Neg. bin.(1)/(2)	Gamma(1)/(2)
TOP10%-share$_{lt}$	−9.884***/***	−11.588***/***	−10.518***/***	−9.809***/***
	(0.215)	(0.343)	(0.309)	(0.294)
log(POP$_{lt}$)*TOP10%-share$_{lt}$	0.720***/***	0.775***/***	0.745***/***	0.724***/***
	(0.015)	(0.022)	(0.020)	(0.019)
log(GDP$_{lt}$)*TOP10%-share$_{lt}$	−0.104***/*	−0.079***/	−0.097***/	−0.108***/
	(0.007)	(0.011)	(0.010)	(0.009)
log(GDP$_{kt}$)	−0.001	−0.011	−0.010	−0.010
	(0.008)	(0.012)	(0.011)	(0.010)
log(GDP$_{lt}$)	0.518***/***	0.555***/***	0.535***/***	0.516***/***
	(0.015)	(0.024)	(0.021)	(0.020)
log(POP$_{kt}$)	0.040**/	0.086***/*	0.083***/*	0.080***/**
	(0.014)	(0.021)	(0.018)	(0.017)
log(POP$_{lt}$)	0.162***/*	0.211***/**	0.221***/**	0.225***/**
	(0.016)	(0.025)	(0.022)	(0.021)
GP-INDEX$_{lt}$	0.041***/***	0.044***/***	0.040***/***	0.037***/**
	(0.002)	(0.003)	(0.003)	(0.002)
PATVAL$_{it}$	0.056***/***	0.079***/***	0.074***/***	0.071***/***
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	−1.368***/***	−1.483***/***	−1.326***/***	−1.212***/***
	(0.054)	(0.079)	(0.070)	(0.066)
COMLANG$_{lt}$	0.006***/	0.015***/*	0.011***/*	0.008***/
	(0.001)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	−0.013***/***	−0.013***/***	−0.014***/***	−0.015***/***
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.003***/	0.006***/	0.005***/	0.004***/
	(0.001)	(0.001)	(0.001)	(0.001)
log(DISSIMI$_{klt}$)	0.006***/***	0.007***/***	0.006***/**	0.005***/**
	(0.001)	(0.001)	(0.001)	(0.001)
log(NOINVENTORS$_i$)	0.013***/**	0.017***/*	0.017***/**	0.017***/***
	(0.001)	(0.001)	(0.001)	(0.001)
log(ARTpc$_{lt}$)	−0.029***/***	−0.045***/***	−0.040***/***	−0.035***/***
	(0.003)	(0.005)	(0.004)	(0.004)
FDI-Inflows$_{kt}$	0.040	0.002	0.025	0.036
	(0.021)	(0.028)	(0.026)	(0.025)
FDI-Inflows$_{lt}$	−0.189***/***	−0.198***/***	−0.189***/***	−0.183***/***
	(0.025)	(0.040)	(0.037)	(0.035)
FDI-Outflows$_{kt}$	−0.026	0.010	−0.015	−0.024
	(0.022)	(0.031)	(0.028)	(0.027)
FDI-Outlows$_{lt}$	0.619***/***	0.755***/***	0.711***/***	0.687***/***
	(0.023)	(0.040)	(0.036)	(0.034)
INVSHARE$_{kt}$	−0.070***/	−0.114***/	−0.094***/	−0.079***/
	(0.019)	(0.030)	(0.026)	(0.024)
INVSHARE$_{lt}$	0.193***/**	0.134***/	0.181***/*	0.209***/**
	(0.019)	(0.029)	(0.026)	(0.025)
log(GDP$_{k(t+1)}$)	0.002	0.007	0.004	0.002
	(0.011)	(0.015)	(0.014)	(0.013)
log(GDP$_{l(t+1)}$)	−0.357***/***	−0.375***/***	−0.368***/***	−0.359***/***
	(0.025)	(0.040)	(0.035)	(0.033)
log(GDP$_{k(t+2)}$)	0.033***/	0.050***/	0.048***/	0.047***/
	(0.009)	(0.013)	(0.012)	(0.011)
log(GDP$_{l(t+2)}$)	−0.198***/**	−0.209***/**	−0.203***/**	−0.197***/**
	(0.015)	(0.024)	(0.021)	(0.020)
χ^2 NACE2 Division	74953.4 ***/***	49139.8 ***/***	55180.8 ***/***	56528.5 ***/***
χ^2 Origin	5791.9 ***/***	4671.0 ***/***	5424.4 ***/***	5749.8 ***/***
χ^2 Destination	100665.5 ***/***	61000.4 ***/***	74426.5 ***/***	79644.7 ***/***
χ^2 Time	18056.7 ***/***	12060.9 ***/***	13668.9 ***/***	14305.7 ***/***
Corr(true,fitted)	0.5386	0.5371	0.5353	0.5340
CV error	0.0863	0.3439	0.3451	0.3460

(1)/(2) indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 5.4. Results extended specification for WID top-5%-share (dep. var. Q^w_{iklt})

	Log-linear$^{(1)}$/$^{(2)}$	Poisson$^{(1)}$/$^{(2)}$	Neg. bin.$^{(1)}$/$^{(2)}$	Gamma$^{(1)}$/$^{(2)}$
TOP5%-share$_{lt}$	−11.178$^{***/***}$	−13.185$^{***/***}$	−11.840$^{***/***}$	−10.973$^{***/***}$
	(0.247)	(0.391)	(0.351)	(0.332)
log(POP$_{lt}$)*TOP5%-share$_{lt}$	0.887$^{***/***}$	0.948$^{***/***}$	0.917$^{***/***}$	0.896$^{***/***}$
	(0.017)	(0.024)	(0.022)	(0.021)
log(GDP$_{lt}$)*TOP5%-share$_{lt}$	−0.169$^{***/**}$	−0.137$^{***/*}$	−0.164$^{***/*}$	−0.180$^{***/*}$
	(0.010)	(0.015)	(0.014)	(0.013)
log(GDP$_{kt}$)	−0.004	−0.016	−0.014	−0.013
	(0.009)	(0.012)	(0.011)	(0.010)
log(GDP$_{lt}$)	0.489$^{***/***}$	0.531$^{***/***}$	0.507$^{***/***}$	0.486$^{***/***}$
	(0.015)	(0.024)	(0.021)	(0.020)
log(POP$_{kt}$)	0.049$^{***/}$	0.101$^{***/*}$	0.096$^{***/**}$	0.092$^{***/**}$
	(0.014)	(0.021)	(0.018)	(0.017)
log(POP$_{lt}$)	0.270$^{***/***}$	0.329$^{***/***}$	0.332$^{***/***}$	0.332$^{***/***}$
	(0.015)	(0.023)	(0.021)	(0.020)
GP-INDEX$_{lt}$	0.045$^{***/***}$	0.050$^{***/***}$	0.045$^{***/***}$	0.042$^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.002)
PATVAL$_{it}$	0.056$^{***/***}$	0.079$^{***/***}$	0.074$^{***/***}$	0.071$^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	−1.375$^{***/***}$	−1.486$^{***/***}$	−1.324$^{***/***}$	−1.206$^{***/***}$
	(0.053)	(0.078)	(0.069)	(0.065)
COMLANG$_{lt}$	0.007$^{***/}$	0.016$^{***/**}$	0.012$^{***/*}$	0.009$^{***/*}$
	(0.001)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	−0.012$^{***/***}$	−0.013$^{***/***}$	−0.015$^{***/***}$	−0.015$^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.002$^{**/}$	0.006$^{***/}$	0.004$^{***/}$	0.003$^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(DISSIMI$_{klt}$)	0.003$^{***/*}$	0.004$^{***/*}$	0.003$^{***/*}$	0.003$^{***/*}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(NOINVENTORS$_i$)	0.013$^{***/**}$	0.017$^{***/**}$	0.017$^{***/**}$	0.017$^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(ARTpc$_{lt}$)	−0.020$^{***/}$	−0.033$^{***/***}$	−0.029$^{***/**}$	−0.025$^{***/*}$
	(0.003)	(0.005)	(0.004)	(0.004)
FDI-Inflows$_{kt}$	0.042$^{*/}$	0.003	0.027	0.039
	(0.021)	(0.028)	(0.026)	(0.025)
FDI-Inflows$_{lt}$	−0.260$^{***/***}$	−0.278$^{***/***}$	−0.272$^{***/***}$	−0.267$^{***/***}$
	(0.024)	(0.040)	(0.037)	(0.035)
FDI-Outflows$_{kt}$	−0.025	0.011	−0.014	−0.023$^{/}$
	(0.022)	(0.031)	(0.028)	(0.027)
FDI-Outlows$_{lt}$	0.698$^{***/***}$	0.835$^{***/***}$	0.797$^{***/***}$	0.774$^{***/***}$
	(0.023)	(0.039)	(0.036)	(0.034)
INVSHARE$_{kt}$	−0.037	−0.074$^{*/}$	−0.059$^{*/}$	−0.047
	(0.019)	(0.030)	(0.026)	(0.024)
INVSHARE$_{lt}$	0.264$^{***/*}$	0.218$^{***/}$	0.265$^{***/*}$	0.292$^{***/**}$
	(0.018)	(0.030)	(0.026)	(0.025)
log(GDP$_{k(t+1)}$)	0.002	0.007	0.004	0.002
	(0.011)	(0.015)	(0.014)	(0.014)
log(GDP$_{l(t+1)}$)	−0.344$^{***/***}$	−0.365$^{***/***}$	−0.353$^{***/***}$	−0.341$^{***/***}$
	(0.025)	(0.040)	(0.035)	(0.033)
log(GDP$_{k(t+2)}$)	0.038$^{***/}$	0.056$^{***/*}$	0.053$^{***/*}$	0.051$^{***/*}$
	(0.009)	(0.013)	(0.012)	(0.011)
log(GDP$_{l(t+2)}$)	−0.206$^{***/**}$	−0.220$^{***/**}$	−0.216$^{***/**}$	−0.211$^{***/*}$
	(0.015)	(0.024)	(0.021)	(0.020)
χ^2 NACE2 Division	74499.2 $^{***/***}$	48701.2 $^{***/***}$	54739.4 $^{***/***}$	56085.0 $^{***/***}$
χ^2 Origin	5841.8 $^{***/***}$	4721.3 $^{***/***}$	5491.5 $^{***/***}$	5823.7 $^{***/***}$
χ^2 Destination	101372.7 $^{***/***}$	61510.5 $^{***/***}$	75089.9 $^{***/***}$	80517.1 $^{***/***}$
χ^2 Time	17762.0 $^{***/***}$	11963.8 $^{***/***}$	13521.1 $^{***/***}$	14131.0 $^{***/***}$
Corr(true,fitted)	0.5391	0.5369	0.5350	0.5337
CV error	0.0859	0.3421	0.3433	0.3442

$^{(1)}$/$^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 5.5. Results extended specification for WID top-1%-share (dep. var. Q^w_{iklt})

	Log-linear$^{(1)}/^{(2)}$	Poisson$^{(1)}/^{(2)}$	Neg. bin.$^{(1)}/^{(2)}$	Gamma$^{(1)}/^{(2)}$
TOP1%-share$_{lt}$	$-15.643^{***/***}$	$-19.004^{***/***}$	$-16.636^{***/***}$	$-15.102^{***/***}$
	(0.374)	(0.579)	(0.514)	(0.485)
log(POP$_{lt}$)*TOP1%-share$_{lt}$	$1.299^{***/***}$	$1.410^{***/***}$	$1.359^{***/***}$	$1.324^{***/***}$
	(0.023)	(0.035)	(0.031)	(0.029)
log(GDP$_{lt}$)*TOP1%-share$_{lt}$	$-0.298^{***/***}$	$-0.251^{***/***}$	$-0.298^{***/***}$	$-0.327^{***/***}$
	(0.019)	(0.028)	(0.025)	(0.023)
log(GDP$_{kt}$)	-0.002	-0.012	-0.010	-0.010
	(0.009)	(0.012)	(0.011)	(0.010)
log(GDP$_{lt}$)	$0.346^{***/***}$	$0.372^{***/***}$	$0.365^{***/***}$	$0.354^{***/***}$
	(0.014)	(0.022)	(0.020)	(0.019)
log(POP$_{kt}$)	$0.044^{**/}$	$0.095^{***/*}$	$0.091^{***/**}$	$0.086^{***/**}$
	(0.014)	(0.021)	(0.018)	(0.017)
log(POP$_{lt}$)	$0.244^{***/***}$	$0.291^{***/***}$	$0.288^{***/***}$	$0.284^{***/***}$
	(0.014)	(0.022)	(0.020)	(0.019)
GP-INDEX$_{lt}$	$0.045^{***/***}$	$0.054^{***/***}$	$0.048^{***/***}$	$0.043^{***/***}$
	(0.002)	(0.003)	(0.002)	(0.002)
PATVAL$_{it}$	$0.056^{***/***}$	$0.079^{***/***}$	$0.074^{***/***}$	$0.071^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-1.196^{***/***}$	$-1.247^{***/***}$	$-1.118^{***/***}$	$-1.025^{***/***}$
	(0.051)	(0.077)	(0.068)	(0.064)
COMLANG$_{lt}$	$0.006^{***/}$	$0.014^{***/*}$	$0.011^{***/*}$	$0.008^{***/}$
	(0.001)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.013^{***/***}$	$-0.014^{***/***}$	$-0.015^{***/***}$	$-0.016^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.002^{***/}$	$0.006^{***/}$	$0.004^{***/}$	$0.004^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(DISSIMI$_{klt}$)	$0.004^{***/**}$	$0.006^{***/**}$	$0.005^{***/**}$	$0.004^{***/*}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(NOINVENTORS$_i$)	$0.013^{***/**}$	$0.017^{***/*}$	$0.017^{***/**}$	$0.017^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)
log(ARTpc$_{lt}$)	$-0.057^{***/***}$	$-0.073^{***/***}$	$-0.066^{***/***}$	$-0.061^{***/***}$
	(0.003)	(0.005)	(0.004)	(0.004)
FDI-Inflows$_{kt}$	0.033	-0.005	0.020	0.031
	(0.021)	(0.028)	(0.026)	(0.025)
FDI-Inflows$_{lt}$	$-0.314^{***/***}$	$-0.318^{***/***}$	$-0.316^{***/***}$	$-0.313^{***/***}$
	(0.024)	(0.040)	(0.037)	(0.035)
FDI-Outflows$_{kt}$	-0.017	0.019	-0.006	-0.016
	(0.022)	(0.031)	(0.028)	(0.027)
FDI-Outlows$_{lt}$	$0.743^{***/***}$	$0.861^{***/***}$	$0.835^{***/***}$	$0.820^{***/***}$
	(0.024)	(0.039)	(0.035)	(0.034)
INVSHARE$_{kt}$	$-0.049^{*/}$	$-0.091^{**/}$	$-0.074^{**/}$	$-0.062^{*/}$
	(0.019)	(0.030)	(0.026)	(0.024)
INVSHARE$_{lt}$	0.013	-0.051	0.005	0.040
	(0.017)	(0.028)	(0.025)	(0.024)
log(GDP$_{k(t+1)}$)	0.004	0.009	0.006	0.004
	(0.011)	(0.015)	(0.014)	(0.013)
log(GDP$_{l(t+1)}$)	$-0.207^{***/**}$	$-0.228^{***/**}$	$-0.221^{***/**}$	$-0.213^{***/**}$
	(0.024)	(0.037)	(0.033)	(0.031)
log(GDP$_{k(t+2)}$)	$0.034^{***/}$	$0.051^{***/}$	$0.048^{***/}$	$0.047^{***/}$
	(0.009)	(0.013)	(0.012)	(0.011)
log(GDP$_{l(t+2)}$)	$-0.164^{***/**}$	$-0.159^{***/*}$	$-0.167^{***/*}$	$-0.169^{***/*}$
	(0.014)	(0.021)	(0.019)	(0.018)
χ^2 NACE2 Division	$76872.1^{***/***}$	$50147.4^{***/***}$	$56422.4^{***/***}$	$57843.0^{***/***}$
χ^2 Origin	$5762.8^{***/***}$	$4688.6^{***/***}$	$5474.1^{***/***}$	$5818.8^{***/***}$
χ^2 Destination	$108616.9^{***/***}$	$64917.8^{***/***}$	$79298.8^{***/***}$	$85269.8^{***/***}$
χ^2 Time	$16235.7^{***/***}$	$11014.9^{***/***}$	$12492.0^{***/***}$	$13077.5^{***/***}$
Corr(true,fitted)	0.5399	0.5384	0.5366	0.5353
CV error	0.0862	0.3440	0.3452	0.3461

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. Every regression includes a constant.

Table 6.1. Results for SWIID Gini pre-tax (alternative dep. var. Q^u_{iklt})

	Log-linear$^{(1)}/^{(2)}$		Poisson$^{(1)}/^{(2)}$		Neg. bin.$^{(1)}/^{(2)}$		Gamma$^{(1)}/^{(2)}$	
GINInet_{lt}	$-11.165^{***/***}$	$-15.055^{***/***}$	$-10.662^{***/***}$	$-14.470^{***/***}$	$-12.492^{***/***}$	$-15.994^{***/***}$	$-12.774^{***/***}$	$-16.088^{***/***}$
	(0.242)	(0.291)	(0.341)	(0.422)	(0.295)	(0.367)	(0.288)	(0.355)
$\log(\text{POP}_{lt})$*GINInet_{lt}	$0.811^{***/***}$	$1.027^{***/***}$	$0.823^{***/***}$	$1.020^{***/***}$	$0.930^{***/***}$	$1.123^{***/***}$	$0.942^{***/***}$	$1.133^{***/***}$
	(0.016)	(0.019)	(0.022)	(0.027)	(0.019)	(0.023)	(0.018)	(0.023)
$\log(\text{GDP}_{lt})$*GINInet_{lt}	$-0.068^{***/***}$	$-0.067^{***/***}$	$-0.090^{***/***}$	$-0.078^{***/***}$	$-0.092^{***/***}$	$-0.089^{***/***}$	$-0.091^{***/***}$	$-0.093^{***/***}$
	(0.004)	(0.004)	(0.005)	(0.006)	(0.004)	(0.005)	(0.004)	(0.005)
$\log(\text{GDP}_{kt})$	$0.037^{***/***}$	-0.024	$0.048^{***/***}$	-0.020	$0.042^{***/***}$	-0.025	$0.038^{***/***}$	$-0.029^{*/}$
	(0.003)	(0.014)	(0.005)	(0.019)	(0.004)	(0.016)	(0.003)	(0.014)
$\log(\text{GDP}_{lt})$	$0.034^{***/***}$	$0.041^{***/**}$	$0.043^{***/***}$	$0.051^{***/***}$	$0.045^{***/***}$	$0.053^{***/***}$	$0.045^{***/***}$	$0.056^{***/***}$
	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
$\log(\text{POP}_{kt})$	$0.340^{***/***}$	$0.300^{***/**}$	$0.438^{***/***}$	$0.385^{***/***}$	$0.436^{***/***}$	$0.373^{***/***}$	$0.427^{***/***}$	$0.363^{***/***}$
	(0.018)	(0.019)	(0.025)	(0.027)	(0.020)	(0.021)	(0.018)	(0.020)
$\log(\text{POP}_{lt})$	$-0.108^{***/***}$	$-0.187^{***/***}$	-0.001	$-0.066^{***/}$	$-0.074^{***/*}$	$-0.143^{***/**}$	$-0.102^{***/**}$	$-0.174^{***/***}$
	(0.011)	(0.012)	(0.015)	(0.018)	(0.013)	(0.016)	(0.013)	(0.015)
GP-INDEX_{lt}	$0.035^{***/***}$	$0.021^{***/***}$	$0.036^{***/***}$	$0.015^{***/***}$	$0.041^{***/***}$	$0.024^{***/***}$	$0.039^{***/***}$	$0.024^{***/***}$
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
PATVAL_{it}	$0.172^{***/***}$	$0.172^{***/***}$	$0.203^{***/***}$	$0.202^{***/***}$	$0.194^{***/***}$	$0.194^{***/***}$	$0.188^{***/***}$	$0.188^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM_{klt}	$-1.365^{***/***}$	$-1.484^{***/***}$	$-1.405^{***/***}$	$-1.590^{***/***}$	$-1.295^{***/***}$	$-1.445^{***/***}$	$-1.217^{***/***}$	$-1.358^{***/***}$
	(0.049)	(0.052)	(0.069)	(0.074)	(0.060)	(0.065)	(0.059)	(0.064)
COMLANG_{lt}	$0.030^{***/***}$	$0.033^{***/***}$	$0.037^{***/***}$	$0.043^{***/***}$	$0.035^{***/***}$	$0.039^{***/***}$	$0.032^{***/***}$	$0.036^{***/***}$
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
$\log(\text{DIST}_{lt})$	$0.008^{***/}$	$0.006^{***/}$	$0.010^{***/}$	$0.008^{***/}$	$0.009^{***/}$	$0.007^{***/}$	$0.008^{***/}$	$0.005^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{TRADE}_{klt})$	$0.045^{***/***}$	$0.040^{***/***}$	$0.048^{***/***}$	$0.042^{***/***}$	$0.050^{***/***}$	$0.044^{***/***}$	$0.050^{***/***}$	$0.044^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls$^{(3)}$		✓		✓		✓		✓
χ^2 NACE2 Division	$78802.7^{***/***}$	$76026.5^{***/***}$	$51841.3^{***/***}$	$49983.4^{***/***}$	$68760.5^{***/***}$	$66437.4^{***/***}$	$72248.5^{***/***}$	$69794.2^{***/***}$
χ^2 Origin	$39200.6^{***/***}$	$25151.7^{***/***}$	$22067.0^{***/***}$	$14143.7^{***/***}$	$28212.3^{***/***}$	$18954.2^{***/***}$	$29595.8^{***/***}$	$20536.0^{***/***}$
χ^2 Destination	$207446.6^{***/***}$	$112687.9^{***/***}$	$124504.5^{***/***}$	$68461.6^{***/***}$	$153923.7^{***/***}$	$83262.1^{***/***}$	$154770.3^{***/***}$	$84344.5^{***/***}$
χ^2 Time	$15514.4^{***/***}$	$11367.9^{***/***}$	$9026.7^{***/***}$	$6912.1^{***/***}$	$12194.3^{***/***}$	$8742.1^{***/***}$	$13028.6^{***/***}$	$9093.2^{***/***}$
Corr(true,fitted)	0.6429	0.6406	0.6406	0.6366	0.6369	0.6331	0.6351	0.6314
CV error	0.1942	0.1942	1.7361	1.7237	1.7509	1.7372	1.7592	1.7448

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. $^{(3)}$ The set of further control variables contains the following variables: $\log(\text{DISSIMI}_{klt})$, $\log(\text{NOINVENTORS}_i)$, $\log(\text{ARTpc}_{lt})$, FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, $\log(\text{GDP}_{k(t+1)})$, $\log(\text{GDP}_{l(t+1)})$, $\log(\text{GDP}_{k(t+2)})$, and $\log(\text{GDP}_{l(t+2)})$. Every regression includes a constant.

TABLE 6.2. Results for SWIID Gini post-tax (alternative dep. var. Q^u_{iklt})

	Log-linear[(1)]/[(2)]		Poisson[(1)]/[(2)]		Neg. bin.[(1)]/[(2)]		Gamma[(1)]/[(2)]	
GINImar$_{lt}$	−6.928***/***	−8.667***/***	−6.540***/***	−8.367***/***	−7.103***/***	−8.672***/***	−7.197***/***	−8.676***/***
	(0.166)	(0.180)	(0.236)	(0.261)	(0.202)	(0.224)	(0.196)	(0.215)
log(POP$_{lt}$)*GINImar$_{lt}$	0.579***/***	0.690***/***	0.605***/***	0.720***/***	0.611***/***	0.714***/***	0.601***/***	0.702***/***
	(0.012)	(0.013)	(0.017)	(0.019)	(0.014)	(0.016)	(0.014)	(0.015)
log(GDP$_{lt}$)*GINImar$_{lt}$	−0.098***/***	−0.104***/***	−0.123***/***	−0.129***/***	−0.110***/***	−0.118***/***	−0.103***/***	−0.113***/***
	(0.003)	(0.003)	(0.004)	(0.005)	(0.003)	(0.004)	(0.003)	(0.004)
log(GDP$_{kt}$)	0.037***/***	−0.023	0.048***/***	−0.018	0.042***/***	−0.023	0.038***/***	−0.027
	(0.003)	(0.014)	(0.005)	(0.019)	(0.004)	(0.016)	(0.003)	(0.014)
log(GDP$_{lt}$)	0.054***/***	0.067***/***	0.067***/***	0.084***/***	0.060***/***	0.075***/***	0.056***/***	0.072***/***
	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)
log(POP$_{kt}$)	0.316***/***	0.260***/**	0.410***/***	0.343***/***	0.403***/***	0.324***/***	0.394***/***	0.313***/**
	(0.018)	(0.019)	(0.025)	(0.027)	(0.020)	(0.021)	(0.018)	(0.020)
log(POP$_{lt}$)	0.080***/***	0.041***/	0.188***/***	0.157***/***	0.153***/***	0.114***/**	0.126***/***	0.080***/
	(0.010)	(0.012)	(0.015)	(0.017)	(0.013)	(0.015)	(0.012)	(0.014)
GP-INDEX$_{lt}$	0.009***/	−0.005***/	0.011***/*	−0.009***/*	0.010***/	−0.004*/	0.009***/	−0.004*/
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
PATVAL$_{it}$	0.172***/***	0.171***/***	0.203***/***	0.202***/***	0.194***/***	0.193***/***	0.188***/***	0.187***/***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	−0.975***/***	−1.088***/***	−0.955***/***	−1.120***/***	−0.933***/***	−1.054***/***	−0.895***/***	−1.002***/***
	(0.050)	(0.053)	(0.070)	(0.076)	(0.062)	(0.066)	(0.060)	(0.064)
COMLANG$_{lt}$	0.030***/***	0.034***/***	0.037***/***	0.043***/***	0.034***/***	0.040***/***	0.032***/***	0.037***/***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
log(DIST$_{lt}$)	0.009***/	0.006***/	0.010***/	0.008***/	0.010***/	0.007***/	0.009***/	0.005***/
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.046***/***	0.041***/***	0.049***/***	0.043***/***	0.052***/***	0.045***/***	0.051***/***	0.045***/***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls[(3)]		✓		✓		✓		✓
χ^2 NACE2 Division	78880.4 ***/***	76098.0 ***/***	51946.8 ***/***	50096.3 ***/***	68804.4 ***/***	66448.9 ***/***	72277.9 ***/***	69783.0 ***/***
χ^2 Origin	37713.6 ***/***	24026.7 ***/***	21320.4 ***/***	13615.6 ***/***	27071.4 ***/***	17959.7 ***/***	28423.3 ***/***	19381.1 ***/***
χ^2 Destination	186178.8 ***/***	104359.2 ***/***	111627.8 ***/***	62371.1 ***/***	139681.5 ***/***	77194.9 ***/***	141736.1 ***/***	78951.8 ***/***
χ^2 Time	16999.1 ***/***	11680.5 ***/***	9947.5 ***/***	7068.2 ***/***	13221.2 ***/***	8899.3 ***/***	13930.1 ***/***	9193.7 ***/***
Corr(true,fitted)	0.6425	0.6402	0.6406	0.6366	0.6369	0.6331	0.6351	0.6314
CV error	0.1943	0.1944	1.7363	1.7239	1.7510	1.7372	1.7593	1.7448

[(1)]/[(2)] indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. [(3)] The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

Table 6.3. Results for WID top-10%-share (alternative dep. var. Q^u_{iklt})

	Log-linear$^{(1)}/^{(2)}$		Poisson$^{(1)}/^{(2)}$		Neg. bin.$^{(1)}/^{(2)}$		Gamma$^{(1)}/^{(2)}$	
TOP10%-share$_{lt}$	$-10.794^{***/***}$	$-8.021^{***/***}$	$-12.285^{***/***}$	$-10.018^{***/***}$	$-11.279^{***/***}$	$-8.656^{***/***}$	$-10.766^{***/***}$	$-8.068^{***/***}$
	(0.259)	(0.299)	(0.449)	(0.480)	(0.352)	(0.401)	(0.328)	(0.380)
log(POP$_{lt}$)*TOP10%-share$_{lt}$	$0.589^{***/***}$	$0.525^{***/***}$	$0.604^{***/***}$	$0.536^{***/***}$	$0.605^{***/***}$	$0.537^{***/***}$	$0.614^{***/***}$	$0.553^{***/***}$
	(0.019)	(0.020)	(0.030)	(0.030)	(0.025)	(0.026)	(0.023)	(0.024)
log(GDP$_{lt}$)*TOP10%-share$_{lt}$	0.016	$-0.040^{***/}$	$0.054^{***/}$	0.018	0.023	$-0.025^{*/}$	0.002	$-0.053^{***/}$
	(0.009)	(0.010)	(0.014)	(0.015)	(0.012)	(0.013)	(0.011)	(0.012)
log(GDP$_{kt}$)	$0.044^{***/***}$	−0.017	$0.057^{***/***}$	−0.022	$0.050^{***/***}$	−0.022	$0.047^{***/***}$	−0.022
	(0.003)	(0.015)	(0.004)	(0.019)	(0.003)	(0.016)	(0.003)	(0.015)
log(GDP$_{lt}$)	$0.036^{***/**}$	$0.653^{***/***}$	$0.016^{**/}$	$0.617^{***/***}$	$0.035^{***/**}$	$0.673^{***/***}$	$0.048^{***/***}$	$0.696^{***/***}$
	(0.004)	(0.021)	(0.006)	(0.034)	(0.005)	(0.028)	(0.005)	(0.026)
log(POP$_{kt}$)	$0.151^{***/*}$	$0.096^{***/}$	$0.170^{***/*}$	$0.115^{***/}$	$0.186^{***/**}$	$0.122^{***/*}$	$0.188^{***/**}$	$0.122^{***/*}$
	(0.020)	(0.021)	(0.028)	(0.030)	(0.023)	(0.025)	(0.021)	(0.023)
log(POP$_{lt}$)	$-0.210^{***/***}$	$-0.106^{***/}$	$-0.024^{/}$	0.017	$-0.147^{***/*}$	$-0.071^{*/}$	$-0.215^{***/***}$	$-0.122^{***/}$
	(0.019)	(0.022)	(0.029)	(0.036)	(0.024)	(0.029)	(0.023)	(0.028)
GP-INDEX$_{lt}$	$-0.006^{*/}$	$-0.014^{***/}$	$-0.006^{/}$	$-0.010^{*/}$	$-0.008^{*/}$	$-0.014^{***/}$	$-0.008^{**/}$	$-0.016^{***/}$
	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
PATVAL$_{it}$	$0.165^{***/***}$	$0.166^{***/***}$	$0.193^{***/***}$	$0.193^{***/***}$	$0.187^{***/***}$	$0.187^{***/***}$	$0.182^{***/***}$	$0.183^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-1.324^{***/***}$	$-1.206^{***/***}$	$-1.627^{***/***}$	$-1.364^{***/***}$	$-1.412^{***/***}$	$-1.224^{***/***}$	$-1.289^{***/***}$	$-1.154^{***/***}$
	(0.076)	(0.084)	(0.108)	(0.117)	(0.097)	(0.105)	(0.095)	(0.103)
COMLANG$_{lt}$	$0.020^{***/***}$	$0.022^{***/***}$	$0.030^{***/***}$	$0.031^{***/***}$	$0.025^{***/***}$	$0.026^{***/***}$	$0.021^{***/***}$	$0.023^{***/***}$
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.006^{***/}$	$-0.012^{***/**}$	$-0.006^{***/}$	$-0.011^{***/**}$	$-0.007^{***/}$	$-0.013^{***/**}$	$-0.008^{***/}$	$-0.014^{***/**}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.035^{***/***}$	$0.029^{***/***}$	$0.035^{***/***}$	$0.030^{***/***}$	$0.038^{***/***}$	$0.033^{***/***}$	$0.039^{***/***}$	$0.034^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls$^{(3)}$		✓		✓		✓		✓
χ^2 NACE2 Division	64301.1 $^{***/***}$	62007.1 $^{***/***}$	42978.8 $^{***/***}$	41341.3 $^{***/***}$	54135.1 $^{***/***}$	52299.3 $^{***/***}$	55951.1 $^{***/***}$	54095.6 $^{***/***}$
χ^2 Origin	22642.0 $^{***/***}$	42047.5 $^{***/***}$	12509.5 $^{***/***}$	21805.9 $^{***/***}$	16252.5 $^{***/***}$	29096.3 $^{***/***}$	17296.8 $^{***/***}$	31281.5 $^{***/***}$
χ^2 Destination	63611.7 $^{***/***}$	54281.3 $^{***/***}$	36557.4 $^{***/***}$	30253.7 $^{***/***}$	48089.4 $^{***/***}$	40085.6 $^{***/***}$	50527.0 $^{***/***}$	42188.7 $^{***/***}$
χ^2 Time	25398.8 $^{***/***}$	21250.5 $^{***/***}$	16665.2 $^{***/***}$	13473.9 $^{***/***}$	20054.6 $^{***/***}$	16614.6 $^{***/***}$	20866.0 $^{***/***}$	17474.7 $^{***/***}$
Corr(true,fitted)	0.6270	0.6283	0.6179	0.6178	0.6149	0.6148	0.6134	0.6134
CV error	0.1848	0.1849	1.4654	1.4778	1.4743	1.4865	1.4792	1.4914

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. $^{(3)}$ The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

Table 6.4. Results for WID top-5%-share (alternative dep. var. Q^{u}_{iklt})

	Log-linear$^{(1)}/^{(2)}$		Poisson$^{(1)}/^{(2)}$		Neg. bin.$^{(1)}/^{(2)}$		Gamma$^{(1)}/^{(2)}$	
TOP5%-share$_{lt}$	$-13.023^{***/***}$	$-8.790^{***/***}$	$-14.880^{***/***}$	$-11.212^{***/***}$	$-13.527^{***/***}$	$-9.413^{***/***}$	$-12.856^{***/***}$	$-8.665^{***/***}$
	(0.302)	(0.350)	(0.502)	(0.542)	(0.399)	(0.455)	(0.374)	(0.434)
log(POP$_{lt}$)*TOP5%-share$_{lt}$	$0.771^{***/***}$	$0.716^{***/***}$	$0.781^{***/***}$	$0.720^{***/***}$	$0.788^{***/***}$	$0.729^{***/***}$	$0.801^{***/***}$	$0.750^{***/***}$
	(0.021)	(0.022)	(0.032)	(0.032)	(0.027)	(0.027)	(0.025)	(0.026)
log(GDP$_{lt}$)*TOP5%-share$_{lt}$	$-0.026^{*/}$	$-0.134^{***/*}$	0.029	$-0.057^{**/}$	-0.017	$-0.119^{***/}$	$-0.046^{**/}$	$-0.156^{***/}$
	(0.013)	(0.014)	(0.019)	(0.021)	(0.016)	(0.018)	(0.016)	(0.017)
log(GDP$_{kt}$)	$0.046^{***/***}$	-0.019	$0.059^{***/***}$	-0.026	$0.052^{***/***}$	-0.025	$0.048^{***/***}$	-0.024
	(0.003)	(0.015)	(0.004)	(0.019)	(0.003)	(0.016)	(0.003)	(0.015)
log(GDP$_{lt}$)	$0.014^{***/}$	$0.636^{***/***}$	-0.005	$0.602^{***/***}$	$0.011^{*/}$	$0.656^{***/***}$	$0.023^{***/}$	$0.676^{***/***}$
	(0.004)	(0.021)	(0.006)	(0.034)	(0.005)	(0.028)	(0.005)	(0.026)
log(POP$_{kt}$)	$0.166^{***/**}$	$0.111^{***/}$	$0.189^{***/**}$	$0.136^{***/}$	$0.204^{***/***}$	$0.140^{***/*}$	$0.205^{***/***}$	$0.139^{***/*}$
	(0.020)	(0.021)	(0.028)	(0.030)	(0.023)	(0.024)	(0.021)	(0.023)
log(POP$_{lt}$)	$-0.159^{***/*}$	-0.009	0.035	$0.125^{***/}$	$-0.096^{***/}$	0.032	$-0.167^{***/**}$	-0.021
	(0.018)	(0.021)	(0.027)	(0.034)	(0.023)	(0.028)	(0.022)	(0.026)
GP-INDEX$_{lt}$	$0.005^{*/}$	$-0.009^{***/}$	0.005	-0.004	0.002	$-0.010^{**/}$	0.002	$-0.012^{***/}$
	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
PATVAL$_{it}$	$0.165^{***/***}$	$0.166^{***/***}$	$0.193^{***/***}$	$0.193^{***/***}$	$0.187^{***/***}$	$0.187^{***/***}$	$0.182^{***/***}$	$0.183^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-1.233^{***/***}$	$-1.160^{***/***}$	$-1.505^{***/***}$	$-1.295^{***/***}$	$-1.281^{***/***}$	$-1.154^{***/***}$	$-1.160^{***/***}$	$-1.086^{***/***}$
	(0.076)	(0.083)	(0.106)	(0.115)	(0.096)	(0.104)	(0.094)	(0.102)
COMLANG$_{lt}$	$0.020^{***/***}$	$0.023^{***/***}$	$0.030^{***/***}$	$0.032^{***/***}$	$0.025^{***/***}$	$0.028^{***/***}$	$0.021^{***/***}$	$0.025^{***/***}$
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.008^{***/}$	$-0.012^{***/**}$	$-0.008^{***/*}$	$-0.012^{***/**}$	$-0.009^{***/}$	$-0.013^{***/**}$	$-0.010^{***/}$	$-0.015^{***/**}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.033^{***/***}$	$0.029^{***/***}$	$0.033^{***/***}$	$0.029^{***/***}$	$0.036^{***/***}$	$0.032^{***/***}$	$0.037^{***/***}$	$0.032^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls$^{(3)}$		✓		✓		✓		✓
χ^2 NACE2 Division	63926.4 $^{***/***}$	61550.5 $^{***/***}$	42840.0 $^{***/***}$	41192.4 $^{***/***}$	53809.2 $^{***/***}$	51961.8 $^{***/***}$	55583.5 $^{***/***}$	53716.7 $^{***/***}$
χ^2 Origin	22431.2 $^{***/***}$	42319.0 $^{***/***}$	12392.3 $^{***/***}$	21974.0 $^{***/***}$	16128.8 $^{***/***}$	29358.9 $^{***/***}$	17171.8 $^{***/***}$	31535.9 $^{***/***}$
χ^2 Destination	62758.0 $^{***/***}$	55155.4 $^{***/***}$	35718.7 $^{***/***}$	30741.1 $^{***/***}$	47209.6 $^{***/***}$	40749.8 $^{***/***}$	49751.7 $^{***/***}$	42942.9 $^{***/***}$
χ^2 Time	24676.2 $^{***/***}$	21193.0 $^{***/***}$	16455.5 $^{***/***}$	13587.2 $^{***/***}$	19741.2 $^{***/***}$	16720.2 $^{***/***}$	20506.4 $^{***/***}$	17548.8 $^{***/***}$
Corr(true,fitted)	0.6270	0.6284	0.6174	0.6173	0.6144	0.6143	0.6129	0.6129
CV error	0.1844	0.1845	1.4603	1.4726	1.4692	1.4814	1.4741	1.4862

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. $^{(3)}$ The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_{i}$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

Table 6.5. Results for WID top-1%-share (alternative dep. var. Q^u_{iklt})

	Log-linear$^{(1)}/^{(2)}$		Poisson$^{(1)}/^{(2)}$		Neg. bin.$^{(1)}/^{(2)}$		Gamma$^{(1)}/^{(2)}$	
TOP1%-share$_{lt}$	$-16.286^{***/***}$	$-11.687^{***/***}$	$-19.196^{***/***}$	$-15.777^{***/***}$	$-16.932^{***/***}$	$-12.683^{***/***}$	$-15.797^{***/***}$	$-11.383^{***/***}$
	(0.482)	(0.536)	(0.750)	(0.802)	(0.618)	(0.683)	(0.587)	(0.653)
log(POP$_{lt}$)*TOP1%-share$_{lt}$	$1.213^{***/***}$	$1.255^{***/***}$	$1.214^{***/***}$	$1.260^{***/***}$	$1.239^{***/***}$	$1.290^{***/***}$	$1.261^{***/***}$	$1.321^{***/***}$
	(0.029)	(0.030)	(0.044)	(0.045)	(0.037)	(0.038)	(0.035)	(0.036)
log(GDP$_{lt}$)*TOP1%-share$_{lt}$	$-0.201^{***/*}$	$-0.389^{***/***}$	$-0.103^{**/}$	$-0.255^{***/**}$	$-0.190^{***/}$	$-0.371^{***/***}$	$-0.239^{***/*}$	$-0.433^{***/***}$
	(0.025)	(0.026)	(0.037)	(0.038)	(0.031)	(0.032)	(0.030)	(0.031)
log(GDP$_{kt}$)	$0.046^{***/***}$	-0.015	$0.059^{***/***}$	-0.020	$0.053^{***/***}$	-0.019	$0.049^{***/***}$	-0.019
	(0.003)	(0.015)	(0.004)	(0.019)	(0.004)	(0.016)	(0.003)	(0.015)
log(GDP$_{lt}$)	$0.015^{***/}$	$0.492^{***/***}$	-0.001	$0.448^{***/***}$	$0.015^{**/}$	$0.513^{***/***}$	$0.025^{***/}$	$0.538^{***/***}$
	(0.004)	(0.020)	(0.006)	(0.032)	(0.005)	(0.026)	(0.005)	(0.025)
log(POP$_{kt}$)	$0.156^{***/*}$	$0.102^{***/}$	$0.183^{***/**}$	$0.130^{***/}$	$0.197^{***/**}$	$0.133^{***/*}$	$0.196^{***/***}$	$0.130^{***/*}$
	(0.020)	(0.021)	(0.028)	(0.030)	(0.023)	(0.025)	(0.021)	(0.023)
log(POP$_{lt}$)	$-0.126^{***/*}$	$-0.112^{***/}$	$0.056^{*/}$	0.010	$-0.092^{***/}$	$-0.095^{***/}$	$-0.167^{***/**}$	$-0.151^{***/*}$
	(0.018)	(0.020)	(0.026)	(0.032)	(0.023)	(0.026)	(0.022)	(0.025)
GP-INDEX$_{lt}$	$0.005^{*/}$	-0.002	$0.009^{*/}$	$0.008^{*/}$	0.006	0.001	0.005	-0.001
	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
PATVAL$_{it}$	$0.165^{***/***}$	$0.166^{***/***}$	$0.193^{***/***}$	$0.193^{***/***}$	$0.187^{***/***}$	$0.187^{***/***}$	$0.182^{***/***}$	$0.183^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-1.117^{***/***}$	$-1.066^{***/***}$	$-1.291^{***/***}$	$-1.137^{***/***}$	$-1.122^{***/***}$	$-1.027^{***/***}$	$-1.025^{***/***}$	$-0.969^{***/***}$
	(0.074)	(0.080)	(0.109)	(0.115)	(0.096)	(0.102)	(0.093)	(0.099)
COMLANG$_{lt}$	$0.019^{***/***}$	$0.021^{***/***}$	$0.029^{***/***}$	$0.030^{***/***}$	$0.024^{***/***}$	$0.025^{***/***}$	$0.020^{***/***}$	$0.023^{***/***}$
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.007^{***/}$	$-0.013^{***/**}$	$-0.008^{***/*}$	$-0.013^{***/**}$	$-0.009^{***/}$	$-0.015^{***/**}$	$-0.010^{***/}$	$-0.016^{***/**}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.034^{***/***}$	$0.029^{***/***}$	$0.034^{***/***}$	$0.029^{***/***}$	$0.037^{***/***}$	$0.032^{***/***}$	$0.039^{***/***}$	$0.032^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls$^{(3)}$		✓		✓		✓		✓
χ^2 NACE2 Division	65319.2 $^{***/***}$	62989.4 $^{***/***}$	43634.1 $^{***/***}$	42012.9 $^{***/***}$	54820.3 $^{***/***}$	53004.3 $^{***/***}$	56609.1 $^{***/***}$	54781.1 $^{***/***}$
χ^2 Origin	22344.7 $^{***/***}$	42920.2 $^{***/***}$	12299.9 $^{***/***}$	22044.6 $^{***/***}$	16118.2 $^{***/***}$	29668.3 $^{***/***}$	17213.5 $^{***/***}$	31953.6 $^{***/***}$
χ^2 Destination	65022.1 $^{***/***}$	59095.5 $^{***/***}$	37015.8 $^{***/***}$	32955.4 $^{***/***}$	48915.3 $^{***/***}$	43564.9 $^{***/***}$	51519.4 $^{***/***}$	45971.1 $^{***/***}$
χ^2 Time	22410.5 $^{***/***}$	19445.9 $^{***/***}$	15155.7 $^{***/***}$	12571.3 $^{***/***}$	18118.6 $^{***/***}$	15408.4 $^{***/***}$	18773.9 $^{***/***}$	16134.0 $^{***/***}$
Corr(true,fitted)	0.6277	0.6290	0.6187	0.6186	0.6157	0.6157	0.6142	0.6142
CV error	0.1848	0.1849	1.4676	1.4800	1.4765	1.4888	1.4816	1.4938

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. $^{(3)}$ The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

TABLE 7.1. Results interaction with patent protection for SWIID Gini post-tax (dep. var. Q^w_{iklt})

	Log-linear(1)/(2)		Poisson(1)/(2)		Neg. bin.(1)/(2)		Gamma(1)/(2)	
GP-INDEX$_{lt}$*GINInet$_{lt}$	0.237***/***	0.183***/***	0.141***/***	0.066***/*	0.159***/***	0.091***/***	0.172***/***	0.108***/***
	(0.009)	(0.010)	(0.010)	(0.012)	(0.009)	(0.011)	(0.009)	(0.010)
GINInet$_{lt}$	−6.369***/***	−9.011***/***	−7.189***/***	−10.713***/***	−8.119***/***	−11.088***/***	−8.472***/***	−11.091***/***
	(0.224)	(0.263)	(0.276)	(0.335)	(0.246)	(0.302)	(0.237)	(0.290)
log(POP$_{lt}$)*GINInet$_{lt}$	0.570***/***	0.676***/***	0.665***/***	0.820***/***	0.716***/***	0.846***/***	0.725***/***	0.842***/***
	(0.014)	(0.017)	(0.018)	(0.021)	(0.016)	(0.019)	(0.015)	(0.018)
log(GDP$_{lt}$)*GINInet$_{lt}$	−0.121***/***	−0.085***/***	−0.136***/***	−0.096***/***	−0.139***/***	−0.103***/***	−0.135***/***	−0.104***/***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
log(GDP$_{kt}$)	0.043***/***	0.048***/***	0.042***/***	0.025*/*	0.040***/***	0.024*/*	0.039***/***	0.024*/*
	(0.002)	(0.009)	(0.003)	(0.012)	(0.003)	(0.010)	(0.003)	(0.009)
log(GDP$_{lt}$)	0.059***/***	0.038***/***	0.066***/***	0.054***/***	0.067***/***	0.052***/***	0.064***/***	0.050***/***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
log(POP$_{kt}$)	−0.606***/***	−0.593***/***	−0.271***/**	−0.268***/**	−0.244***/**	−0.254***/**	−0.235***/**	−0.251***/**
	(0.014)	(0.015)	(0.018)	(0.019)	(0.016)	(0.017)	(0.015)	(0.016)
log(POP$_{lt}$)	−0.051***/	−0.056***/	0.060***/	0.055***/	0.019	0.020	−0.001	0.002
	(0.009)	(0.011)	(0.012)	(0.013)	(0.010)	(0.012)	(0.010)	(0.012)
GP-INDEX$_{lt}$	−0.050***/***	−0.049***/***	−0.011**/	−0.007	−0.014***/	−0.012**/	−0.020***/*	−0.017***/*
	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)
PATVAL	0.057***/***	0.057***/***	0.082***/***	0.082***/***	0.076***/***	0.076***/***	0.073***/***	0.072***/***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	−0.247***/	−0.426***/***	−0.449***/***	−0.641***/***	−0.276***/*	−0.482***/***	−0.173***/	−0.390***/***
	(0.042)	(0.046)	(0.055)	(0.059)	(0.049)	(0.053)	(0.046)	(0.050)
COMLANG$_{lt}$	−0.016***/*	−0.009***/	−0.002	0.008***/	−0.007***/	0.001/	−0.010***/	−0.003**/
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(DIST$_{lt}$)	0.009***/*	0.008***/	0.010***/	0.008***/	0.009***/	0.007***/	0.009***/	0.006***/
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.019***/***	0.014***/***	0.023***/***	0.017***/***	0.022***/***	0.016***/***	0.022***/***	0.015***/***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls(3)		✓		✓		✓		✓
χ^2 NACE2 Division	61011.3 ***/***	58967.4 ***/***	57588.4 ***/***	55632.9 ***/***	62553.0 ***/***	60531.5 ***/***	63800.6 ***/***	61751.8 ***/***
χ^2 Origin	16717.8 ***/***	24440.3 ***/***	9914.4 ***/***	14771.5 ***/***	11186.1 ***/***	17421.9 ***/***	11676.3 ***/***	19009.8 ***/***
χ^2 Destination	255229.0 ***/***	186159.4 ***/***	206202.2 ***/***	140198.9 ***/***	241302.6 ***/***	163800.2 ***/***	254289.1 ***/***	172862.9 ***/***
χ^2 Time	4287.7 ***/***	4614.8 ***/***	3765.0 ***/***	3671.9 ***/***	4624.3 ***/***	4295.2 ***/***	4982.3 ***/***	4545.5 ***/***
Corr(true,fitted)	0.5093	0.5050	0.5486	0.5441	0.5466	0.5421	0.5451	0.5407
CV error	0.1367	0.1374	0.4115	0.4081	0.4130	0.4095	0.4142	0.4106

(1)/(2) indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. (3) The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

Table 7.2. Results interaction with patent protection for SWIID Gini pre-tax (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}$/$^{(2)}$		Poisson$^{(1)}$/$^{(2)}$		Neg. bin.$^{(1)}$/$^{(2)}$		Gamma$^{(1)}$/$^{(2)}$	
GP-INDEX$_{lt}$*GINImar$_{lt}$	0.097$^{***/}$	0.304$^{***/***}$	0.039$^{***/}$	0.210$^{***/***}$	0.018	0.200$^{***/***}$	0.001	0.187$^{***/***}$
	(0.008)	(0.010)	(0.010)	(0.012)	(0.009)	(0.011)	(0.009)	(0.010)
GINImar$_{lt}$	−7.121$^{***/***}$	−8.217$^{***/***}$	−6.740$^{***/***}$	−8.046$^{***/***}$	−7.017$^{***/***}$	−8.092$^{***/***}$	−7.041$^{***/***}$	−8.012$^{***/***}$
	(0.147)	(0.157)	(0.181)	(0.195)	(0.161)	(0.175)	(0.154)	(0.167)
log(POP$_{lt}$)*GINImar$_{lt}$	0.649$^{***/***}$	0.679$^{***/***}$	0.678$^{***/***}$	0.728$^{***/***}$	0.672$^{***/***}$	0.711$^{***/***}$	0.650$^{***/***}$	0.686$^{***/***}$
	(0.010)	(0.011)	(0.013)	(0.014)	(0.011)	(0.012)	(0.011)	(0.012)
log(GDP$_{lt}$)*GINImar$_{lt}$	−0.145$^{***/***}$	−0.144$^{***/***}$	−0.166$^{***/***}$	−0.166$^{***/***}$	−0.153$^{***/***}$	−0.155$^{***/***}$	−0.140$^{***/***}$	−0.144$^{***/***}$
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
log(GDP$_{kt}$)	0.044$^{***/***}$	0.047$^{***/***}$	0.042$^{***/***}$	0.023$^{*/*}$	0.041$^{***/***}$	0.023$^{*/*}$	0.041$^{***/***}$	0.023$^{*/*}$
	(0.002)	(0.009)	(0.003)	(0.012)	(0.003)	(0.010)	(0.003)	(0.009)
log(GDP$_{lt}$)	0.080$^{***/***}$	0.076$^{***/***}$	0.091$^{***/***}$	0.094$^{***/***}$	0.084$^{***/***}$	0.085$^{***/***}$	0.076$^{***/***}$	0.078$^{***/***}$
	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)
log(POP$_{kt}$)	−0.625$^{***/***}$	−0.611$^{***/***}$	−0.292$^{***/***}$	−0.287$^{***/**}$	−0.270$^{***/**}$	−0.280$^{***/**}$	−0.263$^{***/**}$	−0.281$^{***/**}$
	(0.014)	(0.015)	(0.018)	(0.019)	(0.016)	(0.017)	(0.015)	(0.016)
log(POP$_{lt}$)	0.064$^{***/}$	0.102$^{***/**}$	0.173$^{***/***}$	0.206$^{***/***}$	0.161$^{***/**}$	0.194******	0.154$^{***/**}$	0.183******
	(0.009)	(0.010)	(0.011)	(0.013)	(0.010)	(0.011)	(0.009)	(0.011)
GP-INDEX$_{lt}$	−0.020$^{***/}$	−0.140$^{***/***}$	0.010$^{*/}$	−0.099$^{***/***}$	0.018$^{***/}$	−0.091$^{***/***}$	0.024$^{***/}$	−0.084$^{***/***}$
	(0.004)	(0.004)	(0.005)	(0.006)	(0.004)	(0.005)	(0.004)	(0.005)
PATVAL	0.057$^{***/***}$	0.056$^{***/***}$	0.082$^{***/***}$	0.081$^{***/***}$	0.076$^{***/***}$	0.076$^{***/***}$	0.073$^{***/***}$	0.072$^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	0.077	−0.131$^{**/}$	0.039	−0.194$^{***/}$	0.127$^{**/}$	−0.081$^{/}$	0.160$^{***/}$	−0.033$^{/}$
	(0.042)	(0.045)	(0.055)	(0.059)	(0.049)	(0.052)	(0.046)	(0.050)
COMLANG$_{lt}$	−0.015$^{***/*}$	−0.009$^{***/}$	−0.002	0.008$^{***/}$	−0.007$^{***/}$	0.002$^{/}$	−0.010$^{***/}$	−0.003$^{*/}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(DIST$_{lt}$)	0.011$^{***/*}$	0.007$^{***/}$	0.011$^{***/}$	0.007$^{***/}$	0.011$^{***/}$	0.006$^{***/}$	0.010$^{***/*}$	0.006$^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.021$^{***/***}$	0.014$^{***/***}$	0.024$^{***/***}$	0.017$^{***/***}$	0.024$^{***/***}$	0.016$^{***/***}$	0.024$^{***/***}$	0.016$^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls$^{(3)}$		✓		✓		✓		✓
χ^2 NACE2 Division	61002.2 $^{***/***}$	58972.0 $^{***/***}$	57622.8 $^{***/***}$	55685.6 $^{***/***}$	62515.5 $^{***/***}$	60520.2 $^{***/***}$	63751.7 $^{***/***}$	61726.5 $^{***/***}$
χ^2 Origin	16774.3 $^{***/***}$	24071.1 $^{***/***}$	9914.9 $^{***/***}$	14498.5 $^{***/***}$	11115.4 $^{***/***}$	17028.6 $^{***/***}$	11601.6 $^{***/***}$	18562.3 $^{***/***}$
χ^2 Destination	240129.5 $^{***/***}$	177690.4 $^{***/***}$	198537.3 $^{***/***}$	135583.4 $^{***/***}$	234835.1 $^{***/***}$	160638.4 $^{***/***}$	248996.2 $^{***/***}$	170703.7 $^{***/***}$
χ^2 Time	5100.7 $^{***/***}$	4917.9 $^{***/***}$	4469.8 $^{***/***}$	3848.7 $^{***/***}$	5327.7 $^{***/***}$	4416.0 $^{***/***}$	5593.4 $^{***/***}$	4585.1 $^{***/***}$
Corr(true,fitted)	0.5090	0.5051	0.5487	0.5444	0.5467	0.5424	0.5452	0.5410
CV error	0.1367	0.1373	0.4114	0.4079	0.4130	0.4094	0.4141	0.4104

$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. $^{(3)}$ The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

TABLE 7.3. Results interaction with patent protection for WID top-10%-share (dep. var. Q^w_{iklt})

	Log-linear[(1)]/[(2)]		Poisson[(1)]/[(2)]		Neg. bin.[(1)]/[(2)]		Gamma[(1)]/[(2)]	
GP-INDEX$_{lt}$*TOP10%-share$_{lt}$	0.097***/***	0.781***/***	0.937***/***	0.920***/***	0.932***/***	0.922***/***	0.926***/***	0.918***/***
	(0.060)	(0.019)	(0.023)	(0.025)	(0.021)	(0.023)	(0.020)	(0.022)
TOP10%-share$_{lt}$	−7.121***/***	−7.488***/***	−12.213***/***	−9.323***/***	−11.138***/***	−8.123***/***	−10.308***/***	−7.275***/***
	(0.516)	(0.266)	(0.300)	(0.334)	(0.274)	(0.310)	(0.261)	(0.299)
log(POP$_{lt}$)*TOP10%-share$_{lt}$	0.649***/***	0.363***/***	0.406***/***	0.362***/***	0.381***/***	0.340***/***	0.360***/***	0.323***/***
	(0.024)	(0.020)	(0.024)	(0.026)	(0.022)	(0.023)	(0.021)	(0.023)
log(GDP$_{lt}$)*TOP10%-share$_{lt}$	−0.145	−0.089***/	0.024*/	−0.046***/	0.004	−0.073***/*	−0.011	−0.091***/**
	(0.013)	(0.009)	(0.010)	(0.011)	(0.009)	(0.010)	(0.009)	(0.010)
log(GDP$_{kt}$)	0.044***/***	0.068***/*	0.060***/***	0.039***/*	0.059***/***	0.040***/*	0.058***/***	0.040***/***
	(0.005)	(0.009)	(0.003)	(0.011)	(0.003)	(0.010)	(0.002)	(0.009)
log(GDP$_{lt}$)	0.080***/***	0.608***/***	0.123***/***	0.698***/***	0.117***/***	0.660***/***	0.111***/***	0.628***/***
	(0.008)	(0.019)	(0.005)	(0.024)	(0.005)	(0.021)	(0.004)	(0.021)
log(POP$_{kt}$)	−0.625***/***	−0.966***/***	−0.615***/***	−0.679***/***	−0.604***/***	−0.665***/***	−0.599***/***	−0.657***/***
	(0.088)	(0.016)	(0.020)	(0.021)	(0.018)	(0.019)	(0.016)	(0.017)
log(POP$_{lt}$)	0.064***/***	0.056**/	−0.071***/	0.090***/	−0.074***/*	0.092***/	−0.075***/	0.095***/
	(0.043)	(0.019)	(0.020)	(0.024)	(0.019)	(0.022)	(0.018)	(0.021)
GP-INDEX$_{lt}$	−0.020***/***	−0.261***/***	−0.281***/***	−0.287***/***	−0.283***/***	−0.291***/***	−0.284***/***	−0.293***/***
	(0.029)	(0.006)	(0.008)	(0.009)	(0.007)	(0.008)	(0.007)	(0.008)
PATVAL$_{lt}$	0.057***/***	0.053***/***	0.077***/***	0.077***/***	0.072***/***	0.072***/***	0.069***/***	0.069***/***
	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	0.077***/***	−1.042***/***	−1.166***/***	−1.084***/***	−0.991***/***	−0.959***/***	−0.871***/***	−0.868***/***
	(0.126)	(0.057)	(0.073)	(0.079)	(0.065)	(0.071)	(0.061)	(0.067)
COMLANG$_{lt}$	−0.015***/***	−0.026***/	−0.005**/	−0.005**/*	−0.011***/*	−0.011***/**	−0.016***/**	−0.015***/***
	(0.007)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	0.011***/***	−0.018***/***	−0.016***/***	−0.020***/***	−0.017***/***	−0.021***/***	−0.017***/***	−0.022***/***
	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.021**/	−0.003***/	0.002*/	−0.002*/	0.001	−0.003***/	0.001	−0.004***/
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls[(3)]		✓		✓		✓		✓
χ^2 NACE2 Division	53273.7 ***/***	51040.2 ***/***	46399.3 ***/***	44508.7 ***/***	50044.4 ***/***	48139.5 ***/***	50784.6 ***/***	48884.3 ***/***
χ^2 Origin	16154.4 ***/***	13145.4 ***/***	8921.4 ***/***	6761.0 ***/***	10215.5 ***/***	7850.1 ***/***	10770.7 ***/***	8359.8 ***/***
χ^2 Destination	107676.6 ***/***	90902.3 ***/***	76079.2 ***/***	63017.7 ***/***	91228.4 ***/***	75573.9 ***/***	97174.2 ***/***	80471.1 ***/***
χ^2 Time	15149.8 ***/***	12762.5 ***/***	12298.1 ***/***	10665.6 ***/***	13561.6 ***/***	11677.0 ***/***	13944.8 ***/***	11948.6 ***/***
Corr(true,fitted)	0.5024	0.5030	0.5337	0.5337	0.5320	0.5320	0.5306	0.5307
CV error	0.1155	0.1170	0.3307	0.3337	0.3317	0.3347	0.3326	0.3355

[(1)]/[(2)] indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. [(3)] The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

Table 7.4. Results interaction with patent protection for WID top-5%-share (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$		Poisson$^{(1)}/^{(2)}$		Neg. bin.$^{(1)}/^{(2)}$		Gamma$^{(1)}/^{(2)}$	
GP-INDEX$_{lt}$*TOP5%-share$_{lt}$	0.791$^{***/***}$	0.783$^{***/***}$	0.897$^{***/***}$	0.871$^{***/***}$	0.908$^{***/***}$	0.892$^{***/***}$	0.913$^{***/***}$	0.901$^{***/***}$
	(0.018)	(0.025)	(0.031)	(0.034)	(0.029)	(0.031)	(0.027)	(0.029)
TOP5%-share$_{lt}$	−10.530$^{***/***}$	−8.470$^{***/***}$	−15.432$^{***/***}$	−11.266$^{***/***}$	−13.857$^{***/***}$	−9.542$^{***/***}$	−12.706$^{***/***}$	−8.378$^{***/***}$
	(0.232)	(0.306)	(0.341)	(0.382)	(0.309)	(0.353)	(0.294)	(0.339)
log(POP$_{lt}$)*TOP5%-share$_{lt}$	0.372$^{***/***}$	0.588$^{***/***}$	0.667$^{***/***}$	0.614$^{***/***}$	0.634$^{***/***}$	0.583$^{***/***}$	0.610$^{***/***}$	0.563$^{***/***}$
	(0.019)	(0.022)	(0.027)	(0.029)	(0.025)	(0.026)	(0.024)	(0.025)
log(GDP$_{lt}$)*TOP5%-share$_{lt}$	0.007$^{***/}$	−0.203$^{***/**}$	−0.033$^{*/}$	−0.135$^{***/***}$	−0.068$^{***/}$	−0.178$^{***/***}$	−0.092$^{***/}$	−0.206$^{***/***}$
	(0.008)	(0.012)	(0.014)	(0.015)	(0.013)	(0.014)	(0.013)	(0.013)
log(GDP$_{kt}$)	0.060$^{***/***}$	0.064$^{***/}$	0.060$^{***/***}$	0.034$^{**/*}$	0.059$^{***/***}$	0.036$^{***/*}$	0.058$^{***/***}$	0.036$^{***/***}$
	(0.002)	(0.009)	(0.003)	(0.011)	(0.003)	(0.010)	(0.002)	(0.009)
log(GDP$_{lt}$)	0.107$^{***/***}$	0.566$^{***/***}$	0.076$^{***/***}$	0.654$^{***/***}$	0.070$^{***/***}$	0.613$^{***/***}$	0.064$^{***/***}$	0.579$^{***/***}$
	(0.004)	(0.019)	(0.005)	(0.024)	(0.004)	(0.022)	(0.004)	(0.021)
log(POP$_{kt}$)	−0.912$^{***/***}$	−0.947$^{***/***}$	−0.595$^{***/***}$	−0.653$^{***/***}$	−0.583$^{***/***}$	−0.642$^{***/***}$	−0.579$^{***/***}$	−0.635$^{***/***}$
	(0.015)	(0.016)	(0.020)	(0.021)	(0.017)	(0.018)	(0.016)	(0.017)
log(POP$_{lt}$)	−0.167$^{**/}$	0.189$^{***/**}$	0.066$^{***/}$	0.236$^{***/**}$	0.060$^{***/}$	0.234$^{***/**}$	0.056$^{**/}$	0.233$^{***/*}$
	(0.016)	(0.018)	(0.019)	(0.023)	(0.018)	(0.021)	(0.017)	(0.020)
GP-INDEX$_{lt}$	−0.251$^{***/***}$	−0.165$^{***/***}$	−0.151$^{***/***}$	−0.161$^{***/***}$	−0.160$^{***/***}$	−0.171$^{***/***}$	−0.165$^{***/***}$	−0.177$^{***/***}$
	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)
PATVAL	0.053$^{***/***}$	0.053$^{***/***}$	0.077$^{***/***}$	0.077$^{***/***}$	0.072$^{***/***}$	0.072$^{***/***}$	0.069$^{***/***}$	0.069$^{***/***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	−0.956$^{***/***}$	−1.193$^{***/***}$	−1.336$^{***/***}$	−1.276$^{***/***}$	−1.156$^{***/***}$	−1.138$^{***/***}$	−1.032$^{***/***}$	−1.039$^{***/***}$
	(0.052)	(0.056)	(0.073)	(0.078)	(0.065)	(0.070)	(0.061)	(0.066)
COMLANG$_{lt}$	−0.027$^{***/***}$	−0.027$^{***/}$	−0.007$^{***/}$	−0.005$^{**/*}$	−0.013$^{***/*}$	−0.011$^{***/**}$	−0.017$^{***/***}$	−0.015$^{***/***}$
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	−0.013$^{***/***}$	−0.017$^{***/***}$	−0.016$^{***/***}$	−0.019$^{***/***}$	−0.017$^{***/***}$	−0.021$^{***/***}$	−0.018$^{***/***}$	−0.021$^{***/***}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	0.002$^{**/}$	−0.002$^{**/}$	0.002	−0.002	0.001	−0.003$^{**/}$	0.001	−0.003$^{***/}$
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls$^{(3)}$		✓		✓		✓		✓
χ^2 NACE2 Division	53023.1 $^{***/***}$	50778.8 $^{***/***}$	46116.7 $^{***/***}$	44229.2 $^{***/***}$	49717.4 $^{***/***}$	47825.6 $^{***/***}$	50438.9 $^{***/***}$	48555.0 $^{***/***}$
χ^2 Origin	15800.9 $^{***/***}$	13068.3 $^{***/***}$	8862.3 $^{***/***}$	6716.2 $^{***/***}$	10156.3 $^{***/***}$	7806.5 $^{***/***}$	10709.5 $^{***/***}$	8316.7 $^{***/***}$
χ^2 Destination	105607.1 $^{***/***}$	90834.1 $^{***/***}$	74623.0 $^{***/***}$	63401.0 $^{***/***}$	89440.0 $^{***/***}$	75961.8 $^{***/***}$	95316.9 $^{***/***}$	80975.5 $^{***/***}$
χ^2 Time	15428.4 $^{***/***}$	12997.3 $^{***/***}$	12248.9 $^{***/***}$	10725.0 $^{***/***}$	13529.9 $^{***/***}$	11759.0 $^{***/***}$	13954.7 $^{***/***}$	12062.7 $^{***/***}$
Corr(true,fitted)	0.5027	0.5033	0.5334	0.5334	0.5317	0.5317	0.5304	0.5304
CV error	0.1152	0.1167	0.3290	0.3319	0.3300	0.3330	0.3309	0.3338

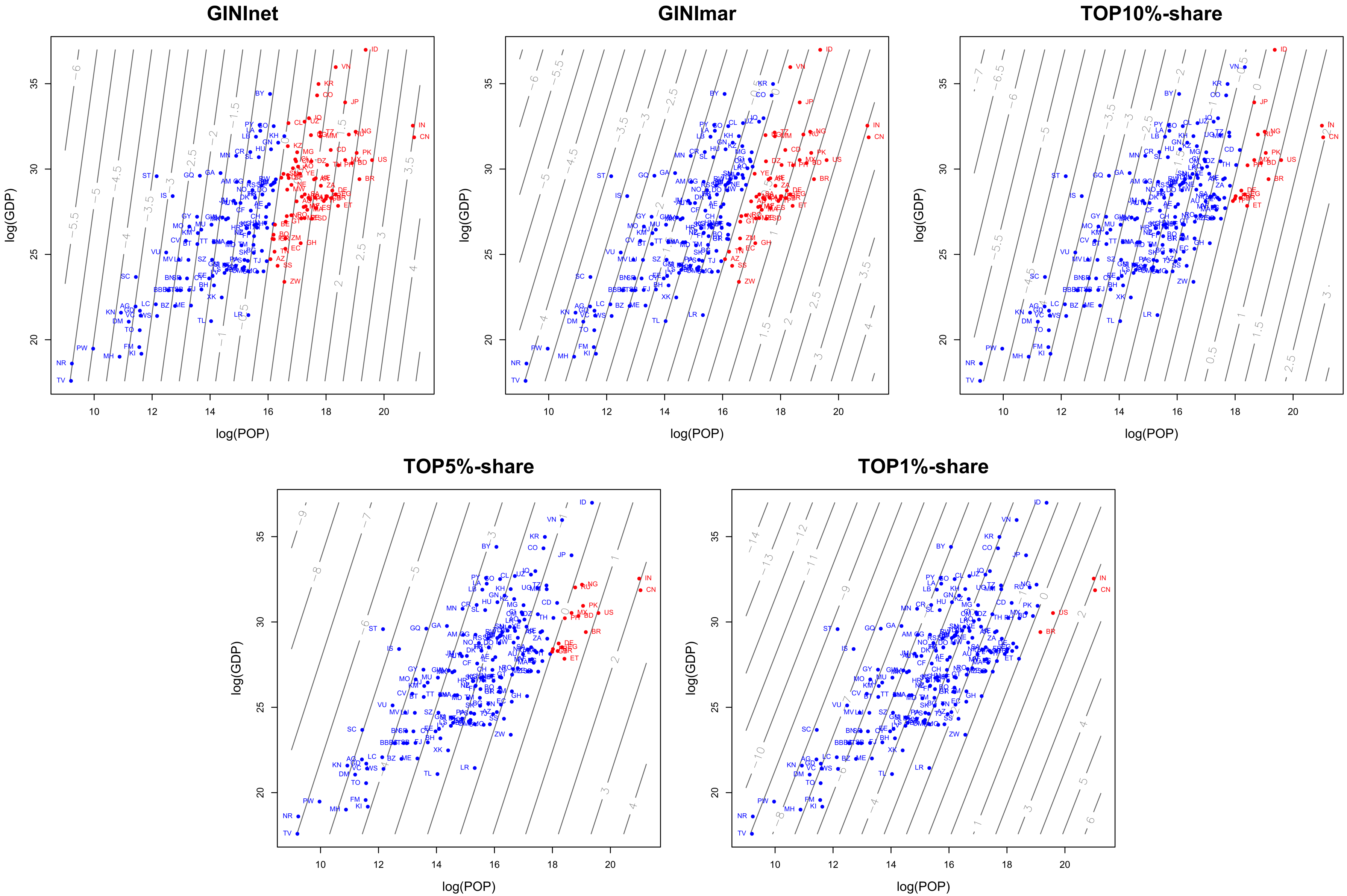
$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively. $^{(3)}$ The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

TABLE 7.5. Results interaction with patent protection for WID top-1%-share (dep. var. Q^{w}_{iklt})

	Log-linear$^{(1)}/^{(2)}$		Poisson$^{(1)}/^{(2)}$		Neg. bin.$^{(1)}/^{(2)}$		Gamma$^{(1)}/^{(2)}$	
GP-INDEX$_{lt}$*TOP1%-share$_{lt}$	$0.791^{***/***}$	$1.094^{***/***}$	$1.007^{***/***}$	$1.078^{***/***}$	$1.045^{***/***}$	$1.141^{***/***}$	$1.067^{***/***}$	$1.175^{***/***}$
	(0.079)	(0.044)	(0.055)	(0.058)	(0.049)	(0.052)	(0.046)	(0.049)
TOP1%-share$_{lt}$	$-10.530^{***/***}$	$-10.483^{***/***}$	$-19.995^{***/***}$	$-15.840^{***/***}$	$-17.424^{***/***}$	$-12.867^{***/***}$	$-15.585^{***/***}$	$-10.869^{***/***}$
	(1.534)	(0.461)	(0.530)	(0.584)	(0.479)	(0.532)	(0.456)	(0.508)
log(POP$_{lt}$)*TOP1%-share$_{lt}$	$0.372^{***/***}$	$1.035^{***/***}$	$1.143^{***/***}$	$1.126^{***/***}$	$1.085^{***/***}$	$1.065^{***/***}$	$1.042^{***/***}$	$1.026^{***/***}$
	(0.060)	(0.031)	(0.040)	(0.041)	(0.036)	(0.037)	(0.033)	(0.035)
log(GDP$_{lt}$)*TOP1%-share$_{lt}$	$0.007^{***/***}$	$-0.487^{***/***}$	$-0.213^{***/**}$	$-0.358^{***/***}$	$-0.267^{***/***}$	$-0.428^{***/***}$	$-0.305^{***/***}$	$-0.475^{***/***}$
	(0.034)	(0.022)	(0.028)	(0.028)	(0.025)	(0.026)	(0.024)	(0.025)
log(GDP$_{kt}$)	$0.060^{***/***}$	$0.066^{***/*}$	$0.059^{***/***}$	$0.037^{***/*}$	$0.058^{***/***}$	$0.038^{***/*}$	$0.057^{***/***}$	$0.039^{***/***}$
	(0.008)	(0.009)	(0.003)	(0.011)	(0.003)	(0.010)	(0.002)	(0.009)
log(GDP$_{lt}$)	$0.107^{***/***}$	$0.382^{***/***}$	$0.046^{***/***}$	$0.449^{***/***}$	$0.040^{***/***}$	$0.425^{***/***}$	$0.036^{***/***}$	$0.404^{***/***}$
	(0.011)	(0.018)	(0.004)	(0.022)	(0.004)	(0.020)	(0.004)	(0.019)
log(POP$_{kt}$)	$-0.912^{***/***}$	$-0.962^{***/***}$	$-0.610^{***/***}$	$-0.665^{***/***}$	$-0.600^{***/***}$	$-0.654^{***/***}$	$-0.596^{***/***}$	$-0.648^{***/***}$
	(0.085)	(0.016)	(0.020)	(0.021)	(0.018)	(0.019)	(0.016)	(0.017)
log(POP$_{lt}$)	$-0.167^{***/***}$	$0.199^{***/***}$	$0.237^{***/***}$	$0.245^{***/***}$	$0.215^{***/***}$	$0.237^{***/***}$	$0.198^{***/***}$	$0.230^{***/**}$
	(0.037)	(0.017)	(0.019)	(0.022)	(0.017)	(0.020)	(0.017)	(0.019)
GP-INDEX$_{lt}$	$-0.251^{***/***}$	$-0.088^{***/***}$	$-0.049^{***/**}$	$-0.059^{***/***}$	$-0.059^{***/**}$	$-0.072^{***/***}$	$-0.066^{***/***}$	$-0.079^{***/***}$
	(0.026)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
PATVAL$_{kt}$	$0.053^{***/***}$	$0.053^{***/***}$	$0.077^{***/***}$	$0.077^{***/***}$	$0.072^{***/***}$	$0.072^{***/***}$	$0.069^{***/***}$	$0.069^{***/***}$
	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TECHSIM$_{klt}$	$-0.956^{***/***}$	$-1.158^{***/***}$	$-1.300^{***/***}$	$-1.219^{***/***}$	$-1.140^{***/***}$	$-1.097^{***/***}$	$-1.028^{***/***}$	$-1.011^{***/***}$
	(0.082)	(0.056)	(0.072)	(0.077)	(0.064)	(0.069)	(0.061)	(0.065)
COMLANG$_{lt}$	$-0.027^{***/***}$	$-0.025^{***/}$	$-0.005^{**/}$	$-0.005^{**/*}$	$-0.011^{***/*}$	$-0.010^{***/**}$	$-0.015^{***/**}$	$-0.014^{***/***}$
	(0.005)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
log(DIST$_{lt}$)	$-0.013^{***/***}$	$-0.016^{***/***}$	$-0.014^{***/***}$	$-0.019^{***/***}$	$-0.015^{***/***}$	$-0.020^{***/***}$	$-0.016^{***/***}$	$-0.020^{***/***}$
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(TRADE$_{klt}$)	$0.002^{***/}$	-0.001	$0.004^{***/}$	-0.001	$0.003^{***/}$	-0.002	$0.003^{***/}$	$-0.002^{*/}$
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls$^{(3)}$		✓		✓		✓		✓
χ^2 NACE2 Division	53823.9 $^{***/***}$	51541.9 $^{***/***}$	47071.6 $^{***/***}$	45160.0 $^{***/***}$	50742.6 $^{***/***}$	48841.5 $^{***/***}$	51486.7 $^{***/***}$	49602.1 $^{***/***}$
χ^2 Origin	15802.3 $^{***/***}$	13001.6 $^{***/***}$	8830.3 $^{***/***}$	6611.8 $^{***/***}$	10203.5 $^{***/***}$	7731.4 $^{***/***}$	10788.0 $^{***/***}$	8273.5 $^{***/***}$
χ^2 Destination	109490.5 $^{***/***}$	93666.1 $^{***/***}$	78317.2 $^{***/***}$	66267.5 $^{***/***}$	93776.5 $^{***/***}$	79702.2 $^{***/***}$	100005.4 $^{***/***}$	85254.4 $^{***/***}$
χ^2 Time	14325.9 $^{***/***}$	12169.5 $^{***/***}$	11255.7 $^{***/***}$	9937.9 $^{***/***}$	12479.8 $^{***/***}$	10990.9 $^{***/***}$	12900.1 $^{***/***}$	11327.9 $^{***/***}$
Corr(true,fitted)	0.5028	0.5033	0.5343	0.5343	0.5326	0.5326	0.5313	0.5313
CV error	0.1159	0.1174	0.3310	0.3340	0.3320	0.3350	0.3329	0.3358

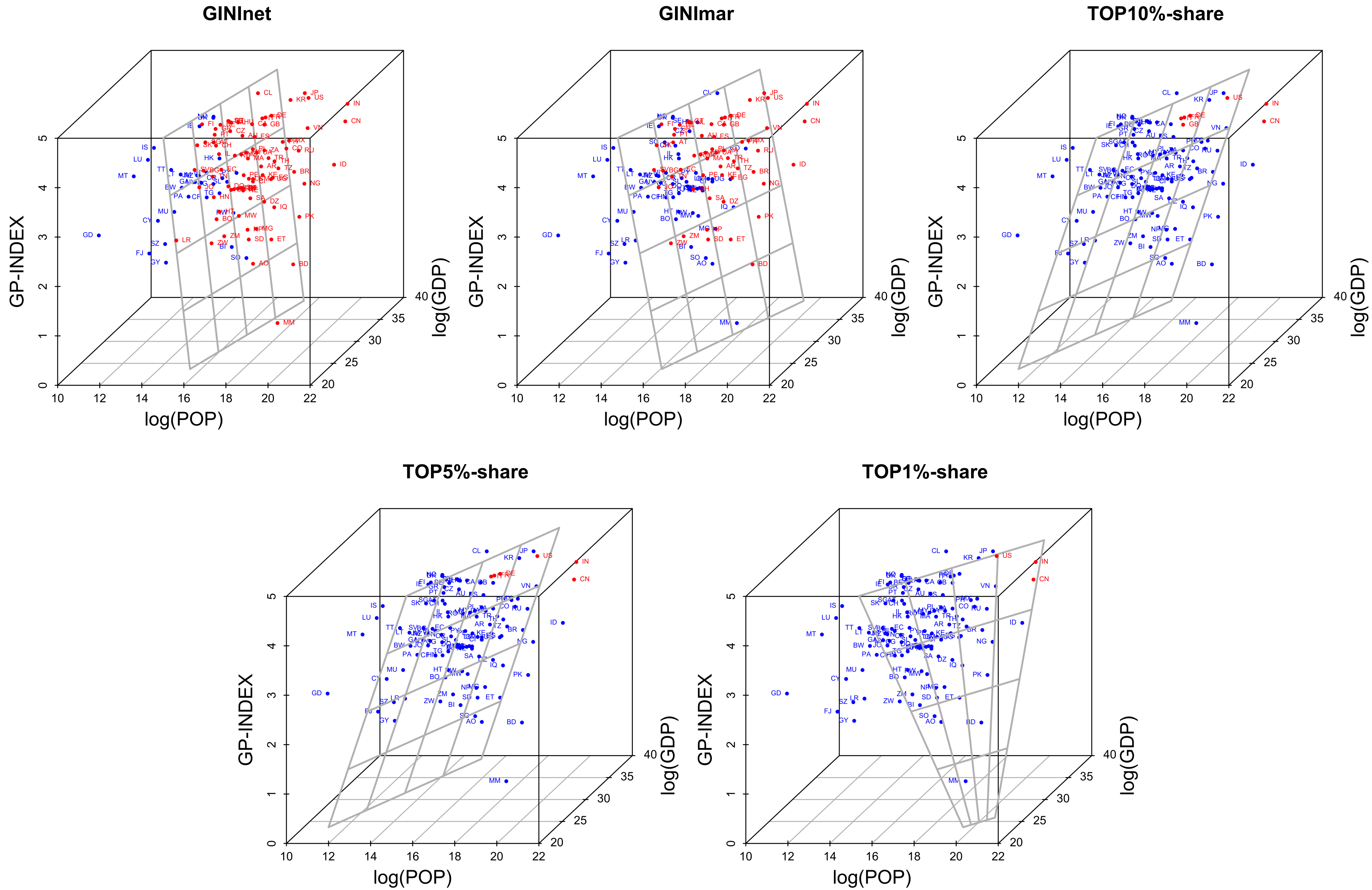
$^{(1)}/^{(2)}$ indicates assumption maintained for inferecne: (1) clustering at the level of the innovation (distinct PATSTAT DOCDB families), and (2) clustering at the level of the NACE2 Division. ***, **, and * indicate significance at 0.01, 0.05, and 0.1, respectively. $^{(3)}$ The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). Every regression includes a constant.

Figure 5. Destination countries, decision boundary, and marginal effects of inequality on innovation



Note: Contour lines for marginal effect of inequality as function of $\log(\text{POP}_{lt})$ and $\log(\text{GDP}_{lt})$ as based on log-linear model estimates for extended set of regressors with dependent variable measured on log scale. Destination-countries are represented by tpopulation and GDP in 2015. Blue (red) points indicate that a reduction in inequality would increase (decrease) innovation.

Figure 6. Destination countries, decision boundary, and sign of marginal effects of inequality on innovation



Note: Countries as points with coordinates log(POP_{lt}), log(GDP_{lt}), and GP-$INDEX_{lt}$ and positive (negative) effect of an increase in inequality in red (blue) as based on log-linear model estimates for extended set of regressors. Destination-countries are represented by population, GDP, and patent protection in 2015. The decision boundary between a positive and a negative marginal effect of an increase in inequality is represented by the indicated gray surface separating the two point-clouds. Blue (red) points indicate that a reduction in inequality would increase (decrease) innovation.

TABLE 8.1. Results control function approach trade for SWIID Gini post-tax (dep. var. Q^w_{iklt})

	Log-linear		Poisson		Neg. bin.		Gamma	
GINInet_{lt}	-8.203^{***}	-10.954^{***}	-8.195^{***}	-11.447^{***}	-9.312^{***}	-12.065^{***}	-9.805^{***}	-12.225^{***}
	(0.251)	(0.665)	(0.289)	(0.817)	(0.263)	(0.734)	(0.254)	(0.700)
$\log(\text{POP}_{lt})*\text{GINInet}_{lt}$	0.688^{***}	0.803^{***}	0.727^{***}	0.868^{***}	0.791^{***}	0.910^{***}	0.810^{***}	0.915^{***}
	(0.015)	(0.057)	(0.019)	(0.068)	(0.017)	(0.065)	(0.016)	(0.063)
$\log(\text{GDP}_{lt})*\text{GINInet}_{lt}$	-0.100^{***}	-0.072^{***}	-0.122^{***}	-0.092^{***}	-0.124^{***}	-0.097^{***}	-0.119^{***}	-0.096^{***}
	(0.003)	(0.015)	(0.004)	(0.017)	(0.004)	(0.017)	(0.003)	(0.017)
$\log(\text{GDP}_{kt})$	0.046^{***}	0.050^{***}	0.043^{***}	0.025^{***}	0.042^{***}	0.025^{***}	0.041^{***}	0.024^{***}
	(0.003)	(0.011)	(0.004)	(0.013)	(0.003)	(0.012)	(0.003)	(0.012)
$\log(\text{GDP}_{lt})$	0.051^{***}	0.044^{***}	0.060^{***}	0.056^{***}	0.061^{***}	0.055^{***}	0.059^{***}	0.054^{***}
	(0.002)	(0.009)	(0.002)	(0.009)	(0.002)	(0.009)	(0.002)	(0.009)
$\log(\text{POP}_{kt})$	-0.623^{***}	-0.603^{***}	-0.281^{***}	-0.274^{***}	-0.256^{***}	-0.262^{***}	-0.248^{***}	-0.259^{***}
	(0.015)	(0.102)	(0.018)	(0.095)	(0.016)	(0.099)	(0.015)	(0.101)
$\log(\text{POP}_{lt})$	-0.010	-0.029	0.091^{***}	0.065	0.049^{***}	0.033	0.029^{***}	0.018
	(0.011)	(0.048)	(0.014)	(0.042)	(0.013)	(0.047)	(0.012)	(0.048)
GP-INDEX_{lt}	0.039^{***}	0.018^{***}	0.045^{***}	0.018^{***}	0.046^{***}	0.022^{***}	0.045^{***}	0.023^{***}
	(0.001)	(0.003)	(0.002)	(0.004)	(0.002)	(0.004)	(0.001)	(0.004)
PATVAL_{it}	0.057^{***}	0.056^{***}	0.082^{***}	0.082^{***}	0.076^{***}	0.076^{***}	0.073^{***}	0.072^{***}
	(0.000)	(0.006)	(0.000)	(0.010)	(0.000)	(0.008)	(0.000)	(0.007)
TECHSIM_{klt}	-0.428^{***}	-0.609^{***}	-0.513^{***}	-0.690^{***}	-0.382^{***}	-0.564^{***}	-0.307^{***}	-0.498^{***}
	(0.050)	(0.120)	(0.065)	(0.110)	(0.058)	(0.105)	(0.056)	(0.103)
COMLANG_{lt}	-0.015^{***}	-0.009^{***}	-0.002	0.008	-0.006^{***}	0.002	-0.010^{***}	-0.003
	(0.001)	(0.007)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)
$\log(\text{DIST}_{lt})$	0.011^{***}	0.008^{***}	0.010^{***}	0.009^{***}	0.010^{***}	0.007^{***}	0.010^{***}	0.007^{***}
	(0.001)	(0.005)	(0.001)	(0.006)	(0.001)	(0.006)	(0.001)	(0.006)
$\log(\text{TRADE}_{klt})$	0.020^{***}	0.014^{***}	0.023^{***}	0.017^{***}	0.023^{***}	0.016^{***}	0.023^{***}	0.015^{***}
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
$\hat{v}_{iklt}$	-0.006^{***}	-0.004^{***}	-0.008^{***}	-0.005^{***}	-0.008^{***}	-0.005^{***}	-0.008^{***}	-0.005^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls included[(1)]		✓		✓		✓		✓
Corr(true,fitted)	0.5091	0.5050	0.5486	0.5441	0.5466	0.5422	0.5451	0.5407
CV error	0.1367	0.1374	0.4115	0.4081	0.4130	0.4095	0.4142	0.4106

Assumption maintained for inference is clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively, for the bootstrap-confidence intervals obtained from a clustered bootstrap over 50 replications. Standard errors reported in parentheses are estimates of the standard deviation of the bootstrap-distribution for the regarding coefficient estimate. Every regression includes a constant and the identified fixed effects δ_j, δ_k, δ_l, and δ_t, respectively. [(1)] The set of further control variables contains the following variables: $\log(\text{DISSIMI}_{klt})$, $\log(\text{NOINVENTORS}_i)$, $\log(\text{ARTpc}_{lt})$, FDI-Inflows_{kt}, FDI-Inflows_{lt}, FDI-Outflows_{kt}, FDI-Outlows_{lt}, INVSHARE_{kt}, INVSHARE_{lt}, $\log(\text{GDP}_{k(t+1)})$, $\log(\text{GDP}_{l(t+1)})$, $\log(\text{GDP}_{k(t+2)})$, and $\log(\text{GDP}_{l(t+2)})$. The linear regression for finding the residuals $\hat{u}_{klt}$ contains the following explanatory variables: $\log(\text{GDPpc}_{kt})$, $\log(\text{GDPpc}_{lt})$, RTA_{klt}, COMCUR_{klt}, SMCTRY_{kl}, COLONY_{kl}, CONTIG_{kl}, $\log(\text{SIZE}_{klt})$, RELFAC_{klt}. For a sample of 23,301 cross-section-times series observations we estimate a linear R^2 of 0.4422. For the regression of $\hat{u}_{klt}$ on all variables included on the right hand side for the main equation we estimate a linear R^2 of 0.8787 when considering all country-pair-time observations where the baseline set of regressors (as shown in the first columns of each panel) is observed, and of 0.8796 for all observations where the the extended set of controls is observed (second columns).

Table 8.2. Results control function approach trade for SWIID Gini pre-tax (dep. var. Q^{w}_{iklt})

	Log-linear		Poisson		Neg. bin.		Gamma	
GINImar_{lt}	−7.066***	−8.222***	−6.712***	−8.188***	−7.000***	−8.148***	−7.037***	−8.010***
	(0.153)	(0.499)	(0.185)	(0.605)	(0.168)	(0.532)	(0.161)	(0.503)
$\log(\text{POP}_{lt})$*GINImar_{lt}	0.653***	0.713***	0.678***	0.761***	0.672***	0.736***	0.651***	0.706***
	(0.011)	(0.027)	(0.013)	(0.042)	(0.012)	(0.035)	(0.011)	(0.032)
$\log(\text{GDP}_{lt})$*GINImar_{lt}	−0.138***	−0.131***	−0.163***	−0.158***	−0.152***	−0.147***	−0.140***	−0.135***
	(0.003)	(0.011)	(0.003)	(0.016)	(0.003)	(0.014)	(0.003)	(0.013)
$\log(\text{GDP}_{kt})$	0.046***	0.050***	0.044***	0.026***	0.042***	0.025***	0.042***	0.025***
	(0.003)	(0.010)	(0.004)	(0.013)	(0.003)	(0.012)	(0.003)	(0.012)
$\log(\text{GDP}_{lt})$	0.077***	0.081***	0.090***	0.098***	0.083***	0.088***	0.076***	0.081***
	(0.001)	(0.009)	(0.002)	(0.010)	(0.002)	(0.009)	(0.002)	(0.008)
$\log(\text{POP}_{kt})$	−0.633***	−0.625***	−0.295***	−0.301***	−0.273***	−0.293***	−0.267***	−0.292***
	(0.015)	(0.103)	(0.018)	(0.095)	(0.016)	(0.099)	(0.015)	(0.101)
$\log(\text{POP}_{lt})$	0.073***	0.095*	0.178***	0.194***	0.163***	0.186***	0.154***	0.179
	(0.009)	(0.047)	(0.013)	(0.041)	(0.011)	(0.043)	(0.011)	(0.043)
GP-INDEX_{lt}	0.023***	0.001***	0.027***	0.001	0.026***	0.002	0.025***	0.003***
	(0.001)	(0.003)	(0.002)	(0.005)	(0.001)	(0.005)	(0.001)	(0.005)
PATVAL_{it}	0.057***	0.056***	0.082***	0.081***	0.076***	0.076***	0.073***	0.072***
	(0.000)	(0.006)	(0.000)	(0.010)	(0.000)	(0.008)	(0.000)	(0.007)
TECHSIM_{klt}	0.103***	−0.042	0.065	−0.081	0.136***	−0.003	0.161***	0.022***
	(0.048)	(0.102)	(0.065)	(0.116)	(0.058)	(0.102)	(0.056)	(0.097)
COMLANG_{lt}	−0.015***	−0.008***	−0.002	0.008	−0.006***	0.002	−0.010***	−0.003
	(0.001)	(0.007)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)
$\log(\text{DIST}_{lt})$	0.011***	0.008***	0.011***	0.008***	0.011***	0.007***	0.010***	0.006***
	(0.001)	(0.004)	(0.001)	(0.006)	(0.001)	(0.006)	(0.001)	(0.005)
$\log(\text{TRADE}_{klt})$	0.021***	0.015***	0.024***	0.017***	0.024***	0.017***	0.024***	0.016***
	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
$\hat{v}_{iklt}$	−0.004***	−0.002*	−0.006***	−0.003*	−0.006***	−0.003***	−0.005***	−0.003***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Further controls included$^{(1)}$		✓		✓		✓		✓
Corr(true,fitted)	0.5090	0.5049	0.5487	0.5443	0.5467	0.5423	0.5452	0.5409
CV error	0.1368	0.1374	0.4114	0.4080	0.4130	0.4094	0.4141	0.4105

Assumption maintained for inference is clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively, for the bootstrap-confidence intervals obtained from a clustered bootstrap over 50 replications. Standard errors reported in parentheses are estimates of the standard deviation of the bootstrap-distribution for the regarding coefficient estimate. Every regression includes a constant and the identified fixed effects δ_j, δ_k, δ_l, and δ_t, respectively. $^{(1)}$ The set of further control variables contains the following variables: $\log(\text{DISSIMI}_{klt})$, $\log(\text{NOINVENTORS}_i)$, $\log(\text{ARTpc}_{lt})$, FDI-Inflows_{kt}, FDI-Inflows_{lt}, FDI-Outflows_{kt}, FDI-Outlows_{lt}, INVSHARE_{kt}, INVSHARE_{lt}, $\log(\text{GDP}_{k(t+1)})$, $\log(\text{GDP}_{l(t+1)})$, $\log(\text{GDP}_{k(t+2)})$, and $\log(\text{GDP}_{l(t+2)})$. The linear regression for finding the residuals $\hat{u}_{klt}$ contains the following explanatory variables: $\log(\text{GDPpc}_{kt})$, $\log(\text{GDPpc}_{lt})$, RTA_{klt}, COMCUR_{klt}, SMCTRY_{kl}, COLONY_{kl}, CONTIG_{kl}, $\log(\text{SIZE}_{klt})$, RELFAC_{klt}. For a sample of 23,301 cross-section-times series observations we estimate a linear R^2 of 0.4422. For the regression of $\hat{u}_{klt}$ on all variables included on the right hand side for the main equation we estimate a linear R^2 of 0.8784 when considering all country-pair-time observations where the baseline set of regressors (as shown in the first columns of each panel) is observed, and of 0.8794 for all observations where the the extended set of controls is observed (second columns).

Table 8.3. Results control function approach trade for WID top-10%-share (dep. var. Q^w_{iklt})

	Log-linear		Poisson		Neg. bin.		Gamma	
TOP10%-share$_{lt}$	−11.776***	−8.941***	−14.045***	−11.298***	−12.697***	−9.904***	−11.684***	−8.926***
	(0.224)	(1.362)	(0.325)	(1.503)	(0.291)	(1.562)	(0.269)	(1.586)
log(POP$_{lt}$)*TOP10%-share$_{lt}$	0.687***	0.690***	0.797***	0.760***	0.758***	0.730***	0.726***	0.708***
	(0.016)	(0.046)	(0.024)	(0.043)	(0.022)	(0.038)	(0.021)	(0.037)
log(GDP$_{lt}$)*TOP10%-share$_{lt}$	−0.027***	−0.122***	−0.016	−0.084***	−0.037***	−0.113***	−0.052***	−0.132***
	(0.008)	(0.043)	(0.010)	(0.048)	(0.009)	(0.051)	(0.009)	(0.054)
log(GDP$_{kt}$)	0.059***	0.064***	0.059***	0.035*	0.057***	0.036*	0.056***	0.036*
	(0.002)	(0.019)	(0.003)	(0.021)	(0.003)	(0.021)	(0.003)	(0.020)
log(GDP$_{lt}$)	0.033***	0.514***	0.038***	0.595***	0.030***	0.551***	0.025***	0.516***
	(0.003)	(0.065)	(0.004)	(0.049)	(0.004)	(0.056)	(0.004)	(0.059)
log(POP$_{kt}$)	−0.914***	−0.962***	−0.610***	−0.669***	−0.601***	−0.658***	−0.599***	−0.652***
	(0.016)	(0.115)	(0.020)	(0.079)	(0.017)	(0.083)	(0.016)	(0.085)
log(POP$_{lt}$)	−0.055***	0.154***	0.046***	0.188***	0.053***	0.202***	0.058***	0.214***
	(0.015)	(0.101)	(0.022)	(0.095)	(0.019)	(0.095)	(0.018)	(0.094)
GP-INDEX$_{lt}$	0.004***	−0.007	0.023***	0.013	0.018***	0.008	0.013***	0.004
	(0.002)	(0.010)	(0.003)	(0.010)	(0.002)	(0.010)	(0.002)	(0.010)
PATVAL$_{it}$	0.054***	0.054***	0.077***	0.077***	0.072***	0.072***	0.069***	0.069***
	(0.000)	(0.006)	(0.000)	(0.009)	(0.000)	(0.007)	(0.000)	(0.007)
TECHSIM$_{klt}$	−1.266***	−1.378***	−1.570***	−1.510***	−1.362***	−1.361***	−1.220***	−1.253***
	(0.048)	(0.075)	(0.075)	(0.104)	(0.068)	(0.091)	(0.064)	(0.084)
COMLANG$_{lt}$	−0.027***	−0.026***	−0.005***	−0.005	−0.011***	−0.010***	−0.015***	−0.014***
	(0.001)	(0.005)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)
log(DIST$_{lt}$)	−0.009***	−0.014***	−0.012***	−0.017***	−0.013***	−0.018***	−0.013***	−0.018***
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
log(TRADE$_{klt}$)	0.005***	0.000	0.005***	0.001	0.005***	0.000	0.005***	0.000
	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.003)
$\hat{v}_{iklt}$	−0.005***	−0.003	−0.009***	−0.007***	−0.008***	−0.006***	−0.007***	−0.006***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Further controls included[(1)]		✓		✓		✓		✓
Corr(true,fitted)	0.5018	0.5024	0.5333	0.5333	0.5316	0.5316	0.5302	0.5303
CV error	0.1156	0.1171	0.3309	0.3339	0.3319	0.3349	0.3328	0.3357

Assumption maintained for inference is clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively, for the bootstrap-confidence intervals obtained from a clustered bootstrap over 50 replications. Standard errors reported in parentheses are estimates of the standard deviation of the bootstrap-distribution for the regarding coefficient estimate. Every regression includes a constant and the identified fixed effects δ_j, δ_k, δ_l, and δ_t, respectively. [(1)] The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). The linear regression for finding the residuals $\hat{u}_{klt}$ contains the following explanatory variables: log(GDPpc$_{kt}$), log(GDPpc$_{lt}$), RTA$_{klt}$, COMCUR$_{klt}$, SMCTRY$_{kl}$, COLONY$_{kl}$, CONTIG$_{kl}$, log(SIZE$_{klt}$), RELFAC$_{klt}$. For a sample of 23,301 cross-section-times series observations we estimate a linear R^2 of 0.4422. For the regression of $\hat{u}_{klt}$ on all variables included on the right hand side for the main equation we estimate a linear R^2 of 0.9018 when considering all country-pair-time observations where the baseline set of regressors (as shown in the first columns of each panel) is observed, and of 0.8989 for all observations where the the extended set of controls is observed (second columns).

TABLE 8.4. Results control function approach trade for WID top-5%-share (dep. var. Q^w_{iklt})

	Log-linear		Poisson		Neg. bin.		Gamma	
TOP5%-share$_{lt}$	-13.531^{***}	-9.379^{***}	-16.362^{***}	-12.363^{***}	-14.624^{***}	-10.591^{***}	-13.366^{***}	-9.394^{***}
	(0.256)	(1.352)	(0.362)	1.514	(0.320)	(1.615)	(0.295)	(1.662)
log(POP$_{lt}$)*TOP5%-share$_{lt}$	0.885^{***}	0.876^{***}	0.990^{***}	0.945^{***}	0.948^{***}	0.915^{***}	0.915^{***}	0.894^{***}
	(0.019)	(0.050)	(0.027)	0.044	(0.024)	(0.040)	(0.023)	(0.041)
log(GDP$_{lt}$)*TOP5%-share$_{lt}$	-0.101^{***}	-0.230^{***}	-0.066^{***}	-0.169^{***}	-0.098^{***}	-0.210^{***}	-0.120^{***}	-0.237^{***}
	(0.011)	(0.055)	(0.014)	0.057	(0.013)	(0.063)	(0.012)	(0.067)
log(GDP$_{kt}$)	0.061^{***}	0.060^{***}	0.060^{***}	0.031^{*}	0.059^{***}	0.032^{*}	0.058^{***}	0.032^{*}
	(0.002)	(0.019)	(0.003)	0.021	(0.003)	(0.021)	(0.003)	(0.020)
log(GDP$_{lt}$)	0.018^{***}	0.491^{***}	0.023^{***}	0.578^{***}	0.015^{***}	0.529^{***}	0.009^{***}	0.490^{***}
	(0.003)	(0.067)	(0.004)	0.050	(0.004)	(0.057)	(0.004)	(0.062)
log(POP$_{kt}$)	-0.901^{***}	-0.951^{***}	-0.595^{***}	-0.652^{***}	-0.587^{***}	-0.644^{***}	-0.585^{***}	-0.639^{***}
	(0.016)	(0.115)	(0.020)	0.079	(0.017)	(0.084)	(0.016)	(0.086)
log(POP$_{lt}$)	0.046^{***}	0.272^{***}	0.150^{***}	0.317^{***}	0.152^{***}	0.324^{***}	0.153^{***}	0.330^{***}
	(0.014)	(0.125)	(0.020)	0.117	(0.018)	(0.118)	(0.017)	(0.117)
GP-INDEX$_{lt}$	0.009^{***}	-0.006	0.031^{***}	0.017^{*}	0.024^{***}	0.011	0.019^{***}	0.006
	(0.002)	(0.009)	(0.003)	0.010	(0.002)	(0.010)	(0.002)	(0.009)
PATVAL$_{it}$	0.053^{***}	0.054^{***}	0.077^{***}	0.077^{***}	0.072^{***}	0.072^{***}	0.069^{***}	0.069^{***}
	(0.000)	(0.006)	(0.000)	0.009	(0.000)	(0.007)	(0.000)	(0.007)
TECHSIM$_{klt}$	-1.251^{***}	-1.381^{***}	-1.539^{***}	-1.518^{***}	-1.331^{***}	-1.361^{***}	-1.189^{***}	-1.249^{***}
	(0.049)	(0.082)	(0.077)	0.097	(0.069)	(0.086)	(0.065)	(0.080)
COMLANG$_{lt}$	-0.028^{***}	-0.026^{***}	-0.006^{***}	-0.004	-0.012^{***}	-0.010^{***}	-0.016^{***}	-0.014^{***}
	(0.001)	(0.005)	(0.002)	0.006	(0.002)	(0.006)	(0.002)	(0.006)
log(DIST$_{lt}$)	-0.011^{***}	-0.015^{***}	-0.014^{***}	-0.017^{***}	-0.014^{***}	-0.018^{***}	-0.015^{***}	-0.018^{***}
	(0.001)	(0.003)	(0.001)	0.003	(0.001)	(0.003)	(0.001)	(0.003)
log(TRADE$_{klt}$)	0.004^{***}	0.000	0.004^{***}	0.000	0.004^{***}	-0.001	0.004^{***}	-0.001
	(0.001)	(0.002)	(0.001)	0.003	(0.001)	(0.003)	(0.001)	(0.003)
$\hat{v}_{iklt}$	-0.005^{***}	-0.003	-0.010^{***}	-0.008^{***}	-0.008^{***}	-0.007^{***}	-0.008^{***}	-0.006^{***}
	(0.001)	(0.001)	(0.001)	0.002	(0.001)	(0.002)	(0.001)	(0.002)
Further controls included$^{(1)}$		✓		✓		✓		✓
Corr(true,fitted)	0.5023	0.5030	0.5333	0.5333	0.5316	0.5316	0.5302	0.5303
CV error	0.1153	0.1168	0.3291	0.3320	0.3301	0.3330	0.3309	0.3338

Assumption maintained for inference is clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively, for the bootstrap-confidence intervals obtained from a clustered bootstrap over 50 replications. Standard errors reported in parentheses are estimates of the standard deviation of the bootstrap-distribution for the regarding coefficient estimate. Every regression includes a constant and the identified fixed effects δ_j, δ_k, δ_l, and δ_t, respectively. $^{(1)}$ The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). The linear regression for finding the residuals $\hat{u}_{klt}$ contains the following explanatory variables: log(GDPpc$_{kt}$), log(GDPpc$_{lt}$), RTA$_{klt}$, COMCUR$_{klt}$, SMCTRY$_{kl}$, COLONY$_{kl}$, CONTIG$_{kl}$, log(SIZE$_{klt}$), RELFAC$_{klt}$. For a sample of 23,301 cross-section-times series observations we estimate a linear R^2 of 0.4422. For the regression of $\hat{u}_{klt}$ on all variables included on the right hand side for the main equation we estimate a linear R^2 of 0.9009 when considering all country-pair-time observations where the baseline set of regressors (as shown in the first columns of each panel) is observed, and of 0.8980 for all observations where the the extended set of controls is observed (second columns).

TABLE 8.5. Results control function approach trade for WID top-1%-share (dep. var. Q^w_{iklt})

	Log-linear		Poisson		Neg. bin.		Gamma	
TOP1%-share$_{lt}$	−15.112***	−11.009***	−19.812***	−16.338***	−17.156***	−13.384***	−15.281***	−11.416***
	(0.393)	(1.648)	(0.525)	1.615	(0.457)	(1.784)	(0.422)	(1.873)
log(POP$_{lt}$)*TOP1%-share$_{lt}$	1.329***	1.352***	1.443***	1.457***	1.377***	1.400***	1.326***	1.360***
	(0.027)	(0.100)	(0.038)	0.079	(0.033)	(0.079)	(0.030)	(0.080)
log(GDP$_{lt}$)*TOP1%-share$_{lt}$	−0.344***	−0.496***	−0.250***	−0.379***	−0.295***	−0.441***	−0.325***	−0.480***
	(0.020)	(0.092)	(0.027)	0.071	(0.023)	(0.079)	(0.021)	(0.084)
log(GDP$_{kt}$)	0.061***	0.063***	0.060***	0.035*	0.058***	0.035*	0.057***	0.035*
	(0.002)	(0.020)	(0.003)	0.022	(0.003)	(0.022)	(0.003)	(0.021)
log(GDP$_{lt}$)	0.021***	0.336***	0.024***	0.408***	0.017***	0.378***	0.013***	0.351***
	(0.003)	(0.066)	(0.004)	0.052	(0.003)	(0.057)	(0.003)	(0.060)
log(POP$_{kt}$)	−0.923***	−0.970***	−0.612***	−0.667***	−0.605***	−0.659***	−0.604***	−0.654***
	(0.016)	(0.120)	(0.020)	0.083	(0.017)	(0.087)	(0.016)	(0.089)
log(POP$_{lt}$)	0.171***	0.243***	0.278***	0.284***	0.259***	0.282***	0.245***	0.281***
	(0.014)	(0.095)	(0.019)	0.090	(0.017)	(0.089)	(0.016)	(0.087)
GP-INDEX$_{lt}$	−0.001	−0.009	0.023***	0.019***	0.016***	0.011	0.010***	0.005
	(0.002)	(0.009)	(0.002)	0.009	(0.002)	(0.009)	(0.002)	(0.009)
PATVAL$_{it}$	0.054***	0.054***	0.077***	0.077***	0.072***	0.072***	0.069***	0.069***
	(0.000)	(0.006)	(0.000)	0.009	(0.000)	(0.007)	(0.000)	(0.007)
TECHSIM$_{klt}$	−1.189***	−1.245***	−1.390***	−1.333***	−1.210***	−1.195***	−1.086***	−1.097***
	(0.054)	(0.080)	(0.077)	0.093	(0.069)	(0.083)	(0.065)	(0.077)
COMLANG$_{lt}$	−0.026***	−0.025***	−0.005***	−0.004	−0.011***	−0.010***	−0.015***	−0.014***
	(0.001)	(0.005)	(0.002)	0.006	(0.002)	(0.006)	(0.002)	(0.006)
log(DIST$_{lt}$)	−0.010***	−0.015***	−0.012***	−0.017***	−0.013***	−0.018***	−0.013***	−0.018***
	(0.001)	(0.002)	(0.001)	0.002	(0.001)	(0.002)	(0.001)	(0.002)
log(TRADE$_{klt}$)	0.006***	0.000	0.005***	0.001	0.005***	0.000	0.005***	0.000
	(0.001)	(0.002)	(0.001)	0.003	(0.001)	(0.003)	(0.001)	(0.003)
$\hat{v}_{iklt}$	−0.003***	−0.002	−0.008***	−0.007***	−0.007***	−0.005***	−0.007***	−0.005***
	(0.001)	(0.001)	(0.001)	0.002	(0.001)	(0.002)	(0.001)	(0.002)
Further controls included(1)		✓		✓		✓		✓
Corr(true,fitted)	0.5026	0.5030	0.5343	0.5342	0.5326	0.5326	0.5313	0.5313
CV error	0.1159	0.1175	0.3310	0.3340	0.3321	0.3350	0.3329	0.3359

Assumption maintained for inference is clustering at the level of the NACE2 Division. ***, ** , and * indicate significance at 0.01, 0.05, and 0.1, respectively, for the bootstrap-confidence intervals obtained from a clustered bootstrap over 50 replications. Standard errors reported in parentheses are estimates of the standard deviation of the bootstrap-distribution for the regarding coefficient estimate. Every regression includes a constant and the identified fixed effects δ_j, δ_k, δ_l, and δ_t, respectively. (1) The set of further control variables contains the following variables: log(DISSIMI$_{klt}$), log(NOINVENTORS$_i$), log(ARTpc$_{lt}$), FDI-Inflows$_{kt}$, FDI-Inflows$_{lt}$, FDI-Outflows$_{kt}$, FDI-Outlows$_{lt}$, INVSHARE$_{kt}$, INVSHARE$_{lt}$, log(GDP$_{k(t+1)}$), log(GDP$_{l(t+1)}$), log(GDP$_{k(t+2)}$), and log(GDP$_{l(t+2)}$). The linear regression for finding the residuals $\hat{u}_{klt}$ contains the following explanatory variables: log(GDPpc$_{kt}$), log(GDPpc$_{lt}$), RTA$_{klt}$, COMCUR$_{klt}$, SMCTRY$_{kl}$, COLONY$_{kl}$, CONTIG$_{kl}$, log(SIZE$_{klt}$), RELFAC$_{klt}$. For a sample of 23,301 cross-section-times series observations we estimate a linear R^2 of 0.4422. For the regression of $\hat{u}_{klt}$ on all variables included on the right hand side for the main equation we estimate a linear R^2 of 0.9021 when considering all country-pair-time observations where the baseline set of regressors (as shown in the first columns of each panel) is observed, and of 0.8993 for all observations where the the extended set of controls is observed (second columns).

Appendix A: Theory

Appendix A1: proof of Proposition 1

Proof. **a)** The left hand side of equation 6 depends negatively on $\widehat{\theta}$, c and N (as $\frac{\partial f(N)}{\partial N}$ falls in N) and positively on Ω and L, while the right hand side rises in F. Consequently, the FE curve in Figure 1 is downward sloping and shifts right when L or Ω increase or when F or c decrease.

As the right hand side of equation 9 increases in $\hat{\theta}$ and Y and decreases in L, F and c, the BC curve is upward sloping and shifts up (left) when Y increases or when L, F and c decrease.

An equilibrium exists if the FE and the BC curve intersect at some point. The BC curve is continuous, goes through the origin and is (weakly) rising. When parameters are such that at least some firms find it profitable to undertake $R\&D$, the FE curve continuously falls until it crosses the $\hat{\theta}$ axis at the positive value $\hat{\theta}_{FE}$ for which $(\Omega - c)\, L(1 - G(\hat{\theta}_{FE})) = F$ holds. As $(1 - G(\hat{\theta}))$ lies between 0 and 1, the FE curve is well defined for all values of N for which $\left(\Omega \frac{\partial f(N)}{\partial N} - c\right) L > F$ holds. As $\frac{\partial f(N)}{\partial N}$ falls in N, the FE curve is therefore well defined for each positive value of N if $\left(\Omega \left(\lim_{N\to\infty} \frac{\partial f(N)}{\partial N}\right) - c\right) L > F$ holds. Under this condition, the FE and the BC curve must therefore cross at some point.

The assumption that θ is distributed within a range $0 \leq \underline{\theta} \leq \theta \leq \bar{\theta}$, where $\underline{\theta} < 1$ is sufficiently low and $\bar{\theta} > 1$ is sufficiently large is made in order to make sure that there are always some households of type $\theta_i < \hat{\theta}$ and others of type $\theta_i > \hat{\theta}$. The assumption that $g(\theta) > 0$ holds in the whole range $\underline{\theta} \leq \theta \leq \bar{\theta}$ is made in order to guarantee that there is always a positive density $g(\hat{\theta})$ of households of type $\theta_i = \hat{\theta}$ purchasing the most expensive innovative good at the limit price (if this was not the case, other equilibria can result in which the price of the most exclusive innovative good lies below the limit price, implying that the free entry condition is not given by equation 6 anymore. In Föllmi and Zweimüller (2006) and Kiedaisch (2016), such "unconstrained" equilibria are analyzed).

b) The distribution of income across housholds with $\theta < \hat{\theta}$ affects the integral on the right hand side of equation 9:$\int_{q=0}^{\hat{\theta}} \frac{L(1-G(q))}{cL(1-G(q))+F} dq \equiv Q$. The integrand $\frac{L(1-G(q))}{cL(1-G(q))+F}$ is positive and a concave falling function of $G(q)$. A regressive transfer among households with $\theta < \hat{\theta}$ implies that the cumulative distribution function $G(q)$ flattens out in the affected region, i.e. that its values are less dispersed. At the same time, such a regressive transfer has no effect on $\int_{q=0}^{\hat{\theta}} G(q)dq = \hat{\theta} G\left(\hat{\theta}\right) - \int_{q=0}^{\hat{\theta}} \theta g(q)dq$ (the expression on the right hand side results from integration by parts) as it leaves the fraction $G\left(\hat{\theta}\right)$ of households with income below $\hat{\theta}$ and the total income share $\int_{q=0}^{\hat{\theta}} \theta g(q)dq$ of these households unaffected. Due to Jensen's inequality, a regressive transfer among households with $\theta < \hat{\theta}$ consequently increases the integral Q and shifts the BC curve up/left. As the transfer does not affect the FE curve, it therefore leads to an increase in N^*.

A transfer from a household with $\theta < \hat{\theta}$ to a household with $\theta > \hat{\theta}$ increases $G(\theta)$ for some $\theta < \hat{\theta}$ (without reducing it for other $\theta < \hat{\theta}$) and therefore reduces the value of the integral Q and shifts the BC curve down/right. As it does not affect the FE curve, it therefore leads to a reduction in N^*.

As neither the BC nor the FE curve depend on the density of households with $\theta > \hat{\theta}$, transfers between these types of households do not affect $\hat{\theta}^*$ or N^*.

c) Given how the different parameters shift the FE and the BC curves (see the proof of part a)), it can easily be seen in Figure 1 that the equilibrium value N^* depends positively on Y and Ω and negatively on F and c. The equilibrium value $\hat{\theta}^*$ increases in Ω and L and decreases in Y. In the case where $\frac{\partial f(C)}{\partial C} = 1\ \forall C$, the FE curve is vertical, implying that $\hat{\theta}^*$ is independent of Y.

When $\hat{\theta}^*$ increases due to an increase in Ω or L or due to a decrease in Y, it increases the fraction $\int_{q=\underline{\theta}}^{\hat{\theta}^*} g(q)dq$ of households that spend all their income on innovative goods and therefore makes it more likely (less likely) that a regressive transfer among two households that are randomly drawn from the distribution $g(\theta)$ affects two households of type $\theta_i < \hat{\theta}^*$ ($\theta_i > \hat{\theta}^*$) and that it therefore increases (decreases) N.

Inserting equation 9 into equation 6 and implicitly differentiating the expression yields $sign\left[\frac{\partial \hat{\theta}}{\partial F}\right] = sign\left[-\Omega\frac{\partial^2 f(N)}{\partial N^2} - \frac{1}{\left(1-G(\hat{\theta})\right)Y\int_{q=0}^{\hat{\theta}}\frac{L(1-G(q))}{(cL(1-G(q))+F)^2}dq}\right]$. As $\frac{\partial^2 f(N)}{\partial N^2} < 0$, $\hat{\theta}$ depends negatively on F when $-\frac{\partial^2 f(N)}{\partial N^2}$ is small, but can depend positively on F when $-\frac{\partial^2 f(N)}{\partial N^2}$ is sufficiently large.

Inserting equation 9 into equation 6 and implicitly differentiating the expression yields $sign\left[\frac{\partial \hat{\theta}}{\partial c}\right] = sign\left[-\Omega\frac{\partial^2 f(N)}{\partial N^2} - \frac{1}{Y\int_{q=0}^{\hat{\theta}}\frac{L(1-G(q))^2}{(cL(1-G(q))+F)^2}dq}\right]$. As $\frac{\partial^2 f(N)}{\partial N^2} < 0$, $\hat{\theta}$ depends negatively on c when $-\frac{\partial^2 f(N)}{\partial N^2}$ is small, but can depend positively on c when $-\frac{\partial^2 f(N)}{\partial N^2}$ is sufficiently large.

In order to show that the effect of L on N can be either positive or negative, the following example is used: suppose that $\frac{\partial f(C)}{\partial C} = 1\ \forall C$, that θ is uniformly distributed and that $\underline{\theta} = 0$. Then, $g(\theta) = \frac{1}{\bar{\theta}}$ and $G(\theta) = \frac{\theta}{\bar{\theta}}$. Equation 6then yields $\hat{\theta} = \left(1 - \frac{F}{L(\Omega-c)}\right)\bar{\theta}$, so that $\frac{\partial \hat{\theta}}{\partial L} = \frac{F\bar{\theta}}{L^2(\Omega-c)}$. Using this and equation 9 allows to derive

$$\frac{\partial N}{\partial L} = \frac{F\bar{\theta}Y\left(1-\frac{\hat{\theta}}{\bar{\theta}}\right)}{L^2(\Omega-c)\left[cL\left(1-\frac{\hat{\theta}}{\bar{\theta}}\right)+F\right]} - Y\int_{s=q}^{\hat{\theta}}\frac{c\left(1-\frac{q}{\bar{\theta}}\right)^2}{\left(cL(1-\frac{q}{\bar{\theta}})+F\right)^2}dq$$

As the first term on the right hand side declines faster than the (absolute value of the) second term when L grows sufficiently large, $\frac{\partial N}{\partial L} < 0$ holds for large values of L. As the integrand on the right hand side is declining in s, $Y\int_{q=0}^{\hat{\theta}} \frac{c\left(1-\frac{q}{\theta}\right)^2}{\left(cL(1-\frac{q}{\theta})+F\right)^2}dq < Y\int_{q=0}^{\hat{\theta}} \frac{c}{(cL+F)^2}dq = Yc\frac{\hat{\theta}}{(cL+F)^2}$ holds. This implies that $\frac{\partial N}{\partial L} > \frac{F\bar{\theta}Y\left(1-\frac{\hat{\theta}}{\bar{\theta}}\right)}{L^2(\Omega-c)\left[cL\left(1-\frac{\hat{\theta}}{\bar{\theta}}\right)+F\right]} - \frac{\hat{\theta}}{(cL+F)^2}$ holds. Inserting $\hat{\theta} = \left(1-\frac{F}{L(\Omega-c)}\right)\bar{\theta}$, the right hand side of this expression is positive when $F\left(cL+F\right)^2 > cL^2\Omega\left(L\left(\Omega-c\right)-F\right)$ holds, which is satisfied if $L = \frac{F}{\Omega-c}$, i.e. if L is set equal to its smallest possible value (below this value, $\hat{\theta}$ would become negative). Therefore, $\frac{\partial N}{\partial L}$ can be either positive or negative. □

Appendix A2: Non-innovative basic need goods

Suppose that there are two types of non-innovative goods: a fixed measure B of "basic need" goods and an infinite variety of producible "backstop" goods. While the latter are identical to the non-innovative goods in the basic model, basic need goods are cheaper to produce and are supplied at the marginal cost of c. They might therefore be interpreted as previously invented innovative goods on which patents have already expired[32]. Households therefore always prefer basic need goods to innovative goods and to backstop goods as the latter two types of goods are sold at prices larger than c in equilibrium. Households with income $y_i < \tilde{y} \equiv Bc$ who are so poor that they cannot afford to purchase one unit of each basic need good (at price c) therefore spend all their income on basic need goods. Households with $y_i > \tilde{y}$ are, however, satiated with basic need goods as they already consume one unit of each of them and therefore also consume other goods. The varieties C_i of innovative goods and G_i of non-innovative goods (including basic need and backstop goods) consumed by household i are therefore given by:

$$C_i = \begin{cases} 0 \quad if & y_i < \tilde{y} \equiv Bc \\ \int_{j=0}^{N} c_{ij}dj \quad if & \tilde{y} < y_i \leq \hat{y} \\ N \quad if & y_i > \hat{y} \end{cases}$$

$$G_i = \begin{cases} \frac{y_i}{c} \quad if & y_i < \tilde{y} \equiv Bc \\ B \quad if & \tilde{y} \leq y_i \leq \hat{y} \\ B + \frac{y_i - Bc - \int_{j=0}^{N} p_j dj}{\Omega} & if \;\; y_i > \hat{y} \end{cases}$$

Let us define $\tilde{\theta} \equiv \frac{\tilde{y}}{\frac{Y}{L}} = \frac{Bc}{\frac{Y}{L}}$, so that households of type $\theta_i \leq \tilde{\theta}$ only consume basic need goods, households of type $\tilde{\theta} < \theta_i < \hat{\theta}$ all basic need goods and some innovative goods and households of type $\theta_i > \hat{\theta}$ all basic need and innovative goods and some backstop goods.

[32] It seems plausible that past innovation efforts have mainly been directed towards basic need goods that have already been in high demand when incomes were lower.

The budget constraint of a household of type $\tilde{\theta} < \theta_i < \hat{\theta}$ is given by

$$\theta_i \frac{Y}{L} = Bc + \int_{q=\tilde{\theta}}^{\theta_i} p(q) dC(q) \tag{15}$$

Proceeding like in the analysis of the basic model (note that $p(q)$ is still given by equation 5) allows to derive the new BC equation as

$$N = \frac{Y}{L} \int_{q=\tilde{\theta}}^{\hat{\theta}} \frac{L(1-G(q))}{cL(1-G(q))+F} dq \tag{16}$$

where $\tilde{\theta} = \frac{BcL}{Y}$. Taking into account that the free entry condition is still given by equation 6, we can derive:

Lemma 1. ***a)*** *A unique equilibrium with positive innovation exists if the conditions from Lemma 1a hold and if B is sufficiently small (precisely, if $\frac{BcL}{Y} < G^{-1}\left(1 - \frac{F}{(\Omega-c)L}\right)$, where $G^{-1}(\bullet)$ is the inverse of $G(\bullet)$).*

b)*: A regressive transfer from a households of type $\theta_i < \hat{\theta}$ to a household of type $\tilde{\theta} \leq \theta_i < \hat{\theta}$ increases N, while a regressive transfer from a household of type $\tilde{\theta} \leq \theta_i \leq \hat{\theta}$ to a household of type $\theta_i > \hat{\theta}$ reduces N. Transfer between households of types $\theta_i < \tilde{\theta}$ and $\theta_i > \hat{\theta}$ do not affect N.*

c)*: $\hat{\theta}$ increases in L and in Ω and decreases in Y. A regressive transfer between two randomly drawn households is more likely to increase and less likely to reduce N the larger Ω is. When $\underline{\theta} \geq \tilde{\theta} = \frac{BcL}{Y}$ (i.e. when all households are rich enough to purchase all basic need goods, but when still $\underline{\theta} < 1$ holds) a regressive transfer between two randomly drawn households is more likely to increase and less likely to reduce N the larger L and the lower Y is. When $\underline{\theta} < \tilde{\theta}$, such a transfer is less likely to reduce N the larger L and the lower Y is, but might not be more likely to increase N. (When $\frac{\partial f(C)}{\partial C} = 1 \ \forall C$ and $\underline{\theta} < \tilde{\theta}$ hold, a regressive transfer is more likely to increase N the larger Y is).*

Proof. **a)** The FE curve remains unchanged. However, the BC curve (equation 16) does not go through the origin now, but leaves the $\hat{\theta}$ axis at the value $\tilde{\theta} = \frac{BcL}{Y}$. The two curves therefore cross if this point lies below $\hat{\theta}_{FE}$, the crossing of the FE curve with the $\hat{\theta}$ axis. As $\hat{\theta}_{FE}$ is defined by the equation $(\Omega - c) L \left(1 - G\left(\hat{\theta}_{FE}\right)\right) = F$, $\tilde{\theta} < \hat{\theta}_{FE}$ holds if $\frac{BcL}{Y} < G^{-1}\left(1 - \frac{F}{(\Omega-c)L}\right)$, where $G^{-1}(\bullet)$ is the inverse of $G(\bullet)$.

b) A transfer from a household of type $\theta_i < \tilde{\theta}$ to a household of type $\tilde{\theta} \leq \theta_i < \hat{\theta}$ decreases $G(\theta)$ for some values in the interval $\tilde{\theta} \leq \theta < \hat{\theta}$ without reducing it for other values in this interval. It therefore shifts the BC curve up/left, leading to an increase in N^*. As neither equation 6 nor equation 16 depend on the density of households with

$\theta_i < \tilde{\theta}$, transfers between these types of households do not affect $\hat{\theta}$ or N. The rest of the proof of part b) is symmetric to the proof of Lemma 1b.

c) As the right hand side of equation 16 increases in $\hat{\theta}$ and Y and decreases in L, the BC curve is upward sloping and shifts up (left) when Y increases or when L decreases. While the analysis is symmetric to the one in the proof of Lemma 1c when $\underline{\theta} \geq \tilde{\theta}$ holds, the additional effect that $\tilde{\theta} = \frac{BcL}{Y}$ increases in L and decreases in Y comes into play in the case where $\underline{\theta} < \tilde{\theta}$ holds. An increse in L and a decline in Y then increase the fraction of households of type $\theta_i < \tilde{\theta}$ among which regressive transfers do not have any effect on N. Keeping $\hat{\theta}$ constant this implies a reduced likelyhood that inequality is good for N. As an increse in L and a decline in Y still lead to an increase in $\hat{\theta}^*$ and therefore make it less likely that a regressive transfer among two randomly selected households is bad for N, whether or not such a transfer is more likely good for N now depends on whether the fraction of people of types $\tilde{\theta} \leq \theta_i < \hat{\theta}$ (among which regressive transfers are good for N) increases or decreases when $\hat{\theta}$ and $\tilde{\theta}$ both increase. This depends on the distribution of θ around the values $\hat{\theta}$ and $\tilde{\theta}$.

When $\frac{\partial f(C)}{\partial C} = 1 \; \forall C$ holds, $\hat{\theta}$ is independent of Y. As $\tilde{\theta}$ falls in Y, a regressive transfer is then more likely to increase N the larger Y is. □

The (qualitative) results therefore stay the same when all households are rich enough to purchase all basic need goods and might only change in a case where sufficiently many households are too poor to afford all basic need goods and in which changes in $\tilde{\theta}$ matter more for the relation between inequality and innovation than changes in $\hat{\theta}$. Applying Lemma 2 to the international context gives the following proposition:

Proposition 4. *Suppose that the conditions from Lemma 1a and the inequality* $\tilde{\theta}_l \equiv \frac{Bc_l L_l}{Y_l} < G^{-1}\left(1 - \frac{F_l}{(\Omega_l - c_l)L_l}\right)$ *are satisfied in a follower country l in which $V_l < N$ holds. Then, the probability ϕ_{jkl} that an innovative good j invented (or first patented) in country k is patented (adopted) in country l depends in the following way on the level of inequality I_l in country l:*

$\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial \Omega_l} > 0$ *always holds and* $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial L_l} > 0$ *and* $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial Y_l} < 0$ *hold when* $\underline{\theta_l} > \tilde{\theta}_l$.

When $\underline{\theta_l} < \tilde{\theta}_l$, $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial L_l} > 0$ *and* $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial Y_l} < 0$ *need not hold for all distributions of* θ_l.

(Suppose that $\frac{\partial f(C)}{\partial C} = 1 \; \forall C$. *Then,* $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial Y_l} = 0$ *holds when* $\underline{\theta_l} > \tilde{\theta}_l$, *but* $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial Y_l} > 0$ *holds when* $\underline{\theta_l} < \tilde{\theta}_l$*).*

Appendix A3: Costless technology transfer

In the basic model, it was assumed that the fixed costs F_l of technology adoption not only include fixed costs of patent application, but also real fixed costs of technology transfer. This implied that no firm had an incentive to adopt a non-patented technology. Let us now instead look at the case where F_l only consists of patent application costs and where there are no real costs of technology adoption. Then, an innovative good that is not patented in country l is imitated and supplied at marginal cost c_l there when at least one household is willing to purchase it. Suppose that $\bar{\theta}_l \frac{Y_l}{L_l} > c_l N$ holds, implying that the income of the richest household in country l is larger than necessary to purchase one unit of each of the globally available N innovative goods at marginal cost c_l. Then, all of the N innovative goods will be produced and supplied in country l, implying that there is always full technology transfer. The level and distribution of income in country l then only determines the measure $V_l \leq N$ of innovative goods that are patented in the country, but not the measure that are produced.

When innovations differred with respect to their cost savings or quality, the more valuable ones would be patented first and an increase in V_l would imply a decrease of the value of the marginal (and average) innovation that gets patented in country l and would therefore imply a decrease in the limit price for this marginal innovation[33]. In order to capture this mechanism without explicitly introducing heterogeneity of goods, the utility function $U_i = f(C_i) + G_i$ (with $\frac{\partial^2 f(C_i)}{\partial C_i^2} < 0$) is now interpreted in a different way: instead of denoting the total measure of innovative goods consumed, C_i now denotes the measure of patent-protected innovative goods consumed by household i, while the consumption of non-patented and non-innovative goods enters the term G_i[34].

In equilibrium, there are again "frontier" countries in which $V_l = N$ holds and follower countries in which $V_l < N$ holds and in which innovators must be indifferent about whether or not to apply for patent protection. Using the notation that households of type $\theta_i \geq \hat{\theta}_l$ are rich enough to consume one unit of each of the N innovative goods

[33] When innovative goods differ with respect to their production costs, it can be shown that those with lower costs end up being sold at a lower price to more households in equilibrium and earn higher profits than those with higher costs. The limit price is therefore indeed binding for the marginal patented innovation that has the highest costs and that is sold at the highest price to the smallest number of households.

[34] As $\frac{\partial^2 f(C)}{\partial C^2} < 0$ and $\left.\frac{\partial f(C)}{\partial C}\right|_{C=0} = 1$ hold, this specification implies that households always value non-patented innovative goods more than patent protected ones, which would not be the case if innovative goods were heterogeneous and if the non-patented ones were of lower quality. This simplifying assumption, however, does not affect the qualitative results of the analysis as long as households prefer to purchase non-patented innovative goods at marginal cost compared to (more valuable) patented ones at the monopoly price.

supplied in a follower country l, the free entry condition is again given by

$$\left(\Omega_l \frac{\partial f(V_l)}{\partial V_l} - c_l\right) L_l(1 - G(\hat{\theta}_l)) = F_l$$

Suppose that all households of type $\theta_i \geq \tilde{\theta}_l$ are rich enough to consume one unit of each of the $N - V_l$ non-patented goods in country l. Proceeding as in Section 2.5, the consumption of patented goods by a household of type θ_i can then be derived as

$$C(\theta_i) = \frac{Y_l}{L_l} \int_{q=\tilde{\theta}}^{\theta_i} \frac{L(1 - G(q))}{cL(1 - G(q)) + F} dq$$

As only households of type $\theta_i > \widehat{\theta}$ consume all patented goods in equilibrium, we can set $C(\hat{\theta}) = V_l$ and derive the BC curve as:

$$V_l = \frac{Y}{L} \int_{q=\tilde{\theta}}^{\hat{\theta}} \frac{L(1 - G(q))}{cL(1 - G(q)) + F} dq \tag{17}$$

As all non-patented goods are sold at marginal cost c_l, the budget constraint of a household of type $\theta_i = \tilde{\theta}_l$ is given by $\tilde{\theta}_l \frac{Y_l}{L_l} = c_l (N - V_l)$, so that $\tilde{\theta}_l = \frac{c_l(N - V_l)L_l}{Y_l}$. The analysis is therefore similar to the one in Appendix B1, except for the fact that there is now an endogenous measure $N - V_l$ of non-patented innovative goods instead of an exogenously given measure B_l of basic need goods.

Proposition 5. *Suppose that all households in country l are rich enough to purchase one unit of each of the non-patented innovative goods. Then, an equilibrium in which $V_l < N$ holds exists in a follower country l if $\left(\Omega_l \left(\frac{\partial f(N)}{\partial N}\right) - c_l\right) L_l > F_l$, if $G^{-1}\left(1 - \frac{F_l}{(\Omega_L - c_l)L_l}\right) > \frac{c_l N L_l}{Y_l}$ (and if the left and right hand side of the inequality are sufficiently close to each other), if $Y_l > \frac{c_l L_l}{c_l L_l + F_l}$ and if $\frac{c_l\left(N - V_l^*\right)L_l}{Y_l} < 1$ hold (the latter inequality always holds if $\frac{c_l N L_l}{Y_l} < 1$).*

The probability ϕ_{jkl} that an innovative good j invented (or first patented) in country k is patented in country l then depends in the following way on the level of inequality I_l in country l: $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial \Omega_l} > 0$, $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial L_l} > 0$ and $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial Y_l} < 0$ (when $\frac{\partial f(C)}{\partial C} = 1\ \forall C$, $\frac{\partial E\left(\frac{\partial \phi_{jkl}}{\partial I_l}\right)}{\partial Y_l} = 0$).

Proof. Proceeding like in the proof of Lemma 1a, it can be shown that the FE curve is well defined for each value $V_l < N$ when $\left(\Omega_l \left(\frac{\partial f(N)}{\partial N}\right) - c_l\right) L_l > F_l$ holds and that it crosses the $\hat{\theta}$ axis at the value $\hat{\theta}_{FE} = G^{-1}\left(1 - \frac{F_l}{(\Omega_L - c_l)L_l}\right)$. The BC curve crosses the $\hat{\theta}$ axis at the value $\hat{\theta}_{BC} = \frac{c_l N L_l}{Y_l}$. When all households are rich enough to purchase one unit of each of the non-patented goods, $G(\tilde{\theta}_l) = 0$ holds. Under this condition, the BC curve is upward sloping (in $\hat{\theta}$ - V_l space) when $Y_l > \frac{c_l L_l}{c_l L_l + F_l}$ holds and downward sloping when $Y_l < \frac{c_l L_l}{c_l L_l + F_l}$ holds.

As households of type $\theta_i > \hat{\theta}^*$ must be rich enough to purchase one unit of each of the N goods at marginal cost, $\hat{\theta}^* \frac{Y_l}{L_l} > c_l N$ must hold, implying that $\hat{\theta}^* > \hat{\theta}_{BC}$ must hold. The only possible equilibrium is therefore the one in which $Y_l > \frac{c_l L_l}{c_l L_l + F_l}$ holds so that the BC curve is upward sloping and in which$\hat{\theta}_{FE} > \hat{\theta}_{BC}$ holds. An equilibrium then exists if the FE and the BC curve cross at a value $V_l^* < N$. As the FE curve is downward sloping (or vertical) this always happens as long as $\hat{\theta}_{BC}$ and $\hat{\theta}_{FE}$ are sufficiently close to each other. Moreover, $\tilde{\theta}_l = \frac{c_l \left(N - V_l^*\right) L_l}{Y_l} < 1$ must hold to guarantee that $\int_{\underline{\theta}}^{\bar{\theta}} \theta g(\theta) d\theta = 1$ can hold.

As the BC curve shifts right when L increases or when Y decreases and as the FE curve shifts right when L or Ω increase and when Y decreases (and does not depend on Y when $\frac{\partial f(C)}{\partial C} = 1 \; \forall C$ holds), the equilibrium value $\hat{\theta}^*$ again depends positively on L and Ω and negatively on Y (or not at all on Y when $\frac{\partial f(C)}{\partial C} = 1 \; \forall C$ holds). Proceeding like in the proof of Lemma 1b, it can again be shown that a regressive transfer among households of type $\theta_i < \hat{\theta}$ rotates the BC curve left while a regressive transfer from households of type $\theta_i < \hat{\theta}$ (who only consume innovative goods) to households of type $\theta_i > \hat{\theta}$ (who consume all N innovative goods and also some non-innovative goods) rotates the BC curve right. Combing these results and using the same arguments used for Proposition 2 then gives the results. □

As long as all (or most) households are rich enough to purchase one unit of each of the non-patented goods, the results from Propostion 2 are therefore robust in this alternative setting[35] . The way in which patent applications depend on the extent of inequality in a country therefore does not seem to depend much on specific assumptions about the transfer of technology and should be indicative of the qualitative effects that inequality has on innovation.

The feature that the transfer of technology to a country does not depend on whether or not an innovator patents an innovation in this country changes the welfare effects of different policies: unlike in the case where a patent application goes along with a transfer of technology, countries in which $V_l < N$ holds can for example never gain from enforcing patents in this context. The reason for that is that patents only lead to a rise in prices and to a reduction in consumption without increasing the world technology frontier N (as innovators do not earn any profits in excess of the patent application fees F_l when $V_l < N$ holds).

[35]When some households are so poor that they only purchase non-patented goods, issues similar to those analyzed in Appendix B1 arise, implying that the results might then change for certain distributions of θ.

Appendix B: Empirical Analysis

Appendix B1: Further Descriptives

Detailed country-lists

Table B1 contains a complete list of all countries surfacing in the SWIID- and WDI-samples samples as origins or destinations of patent-flows. Moreover it contains a list of the 184 (115) countries used for the predictive exercises in the main text.

[Table B1]

Correlation among inequality measures, population, and GDP

Table B2 summarizes all distinct pair-wise correlations among the inequality measures GINInet_{it}, GINImar_{it}, TOP10%-share$_{lt}$, TOP5%-share$_{lt}$, and TOP1%-share$_{lt}$, respectively, and the key variables $\log(\text{GDP}_{lt})$ and $\log(\text{POP}_{lt})$; we have also added a measure of $\log(\text{GDPpc}_{lt}) = \log(\text{GDP}_{lt})$ - $\log(\text{POP}_{lt})$. All inequality measures are highly positively correlated. Except for TOP1%-share$_{lt}$ we would conclude that a large population is positively correlated with inequality. Further we find that within our sample a high Gini coefficient (pre- and post-taxes) is negatively related to a country's GDP per capita, i.e. inequality is likely to be higher for poorer countries in our sample.

[Table B2]

Appendix B2: Construction of Patent Value Variable

As referred to above, we construct a measure of patent-value PATVAL_{it} from five distinct measures of the spreading out of subsequent applications from a patent-family i and as constructed from the raw-data extracted from PATSTAT: (1) A variable counting the number of previous subsequent applications for the family up to time t (PREVIOUS_{it}), (2) a variable counting the number of contemporaneous subsequent applications at time t (CURRENT_{it}), (3) a variable using (1) normalized by the innovation's age where minimum age is set to one (AVSUB_{it}), (4) a variable counting the number of different destination countries the family has already been filed at in t including the current (NODESTINS_{it}), and finally (5) a variable counting the the number of distinct countries where patent-protection is applied for simultaneously for any member of the family in a year t ($\text{NOSIMULTDESTINS}_{it}$). We illustrate this here for the DOCDB-family definition.

Since all of these variables have strictly positive support, we log-transform each of them. All of these measures are highly positively correlated as can be seen from Table B3 with pairwise correlations ranging from 0.577 to 0.957.

To reduce the number of measures, we extract the first principal component of the correlation matrix in Table B3. The according eigenvalue accounts for a share of 79.86% of total variation. The variable loadings of the original five variables on the the variable PATVAL_{it} are shown in Table B4. As is convenient for interpretability, the results indicate a positive relationship among the patent-value variable and the original five components.

Appendix B3: Missing Value Imputation for WDI-Series

To ensure the widest possible coverage of simultaneously observed instances on the dependent variable and a measure of inequality by data on other explanatory variables, we pre-process the raw-data extracted from the WDI with respect to missing values before defining the regressors used above. In specific, the pattern of missingness faced here is one mainly due to gaps in the time series.

Besides total population (SP.POP.TOTL) which is always observed when required, these raw-variables include real GDP in local currency units (NY.GDP.MKTP.CN), net-inflows and net-outflows of foreign direct investment as a share of GDP (BX.KLT. DINV.WD.GD.ZS and BM.KLT.DINV. WD.GD.ZS), the number of technical and scientific journal articles (IP.JRN. ARTC.SC), and capital formation as a share of GDP (NE.GDI.TOTL.ZS) with the regarding WDI-internal names given in parentheses.

For these series a share of 59.97% of the raw-data is completely observed, for 18.93% one variable has a missing entry, for 7.11% two, for 4.91% three, for 4.20% four, and for 5.34% five out of six variables have a missing entry.

For imputation we apply the missing forests algorithm suggested in Bühlman and Stekhoven (2012). Their multiple imputation algorithm is based on iterative prediction of the missing data form random forest models on the observed part of the data. The pattern of missingness assumed here is that missingness does not depend on unobservables (missingness at random). In order to exploit information in the panel structure of the data, we run three steps of imputation: Firstly, we employ missing forests imputation country-wise exploiting the time series variation for each country, then we apply missing forests imputation for each point in time exploiting cross-sectional variation. In a third step, adding the imputed values from these two exercises as additional predictors of missingness to the original data, the missing forests algorithm is applied to impute the missing values in the original data for all observations simultaneously. This is a practical approach which yields a better quality of imputation for the data at hands. Each missing forests imputation relies on 1,500 trees.

Figure B1 plots the densities of the indicated original five variables exhibiting missingness versus their imputed versions. Computing the normalized root-mean-squared error, measuring the quality of the imputation, we obtain a value of approx. 0.080 indicating good quality of the imputation. Moreover, we replicate the main results presented in this paper based on the original WDI-series and do not find namable sensitivities. These additional results can be found in the online appendix to this paper.

Appendix B4: Further Estimation Results

[TO BE ADDED]

Table B1. Country coverage

Country	ISO2	SWIID-samples BASELINE All	Orig.	Dest.	SWIID-samples EXTENDED All	Orig.	Dest.	WID-samples BASELINE All	Orig.	Dest.	WID-samples EXTENDED All	Orig.	Dest.	BASELINE All	Orig.	Dest.	EXTENDED All	Orig.	Dest.	BASELINE All	Orig.	Dest.	EXTENDED All	Orig.	Dest.
Argentina	AR	x	x	x	x		x	x	x					x	x	x	x		x	x	x	x	x		x
Australia	AU	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Austria	AT	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Belgium	BE	x	x	x	x	x	x	x	x					x	x					x	x				
Brazil	BR	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Bulgaria	BG	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Canada	CA	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Chile	CL	x		x	x		x																		
China	CN	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Croatia	HR	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Czech Republik	CZ	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Denmark	DK	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Egypt	EG	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Estonia	EE	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Finland	FI	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
France	FR	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Georgia	GE	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Germany	DE	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Greece	GR	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Hong Kong	HK	x		x																					
Hungary	HU	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Iceland	IS	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Indonesia	ID	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x	x	x	x	x
Ireland	IE	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x		x	x	x	x	x	x
Italy	IT	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Japan	JP	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Jordan	JO	x		x	x		x																		
Kenya	KE	x		x	x		x																		
Korea	KR	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Latvia	LV	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Lithuania	LT	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Luxembourg	LU	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Malawi	MW	x	x	x	x		x																		
Malaysia	MY	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Malta	MT	x	x		x	x																			

Table continued on next page.

Country	ISO2	SWIID-samples						WID-samples																	
		BASELINE			EXTENDED			BASELINE			EXTENDED			BASELINE			EXTENDED			BASELINE			EXTENDED		
		All	Orig.	Dest.	All	Orig.	Dest.	All	Orig.	Dest.	All	Orig.	Dest.	All	Orig.	Dest.	All	Orig.	Dest.	All	Orig.	Dest.	All	Orig.	Dest.
Mexico	MX	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Morocco	MA	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Netherlands	NL	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
New Zealand	NZ	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Nicaragua	NI	x		x	x		x																		
Norway	NO	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Philippines	PH	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Poland	PL	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Portugal	PT	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Romania	RO	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Russia	RU	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Singapore	SG	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x		x
Slovakia	SK	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
Slovenia	SI	x	x		x	x		x	x		x	x		x	x		x	x		x	x		x	x	
South Africa	ZA	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
Spain	ES	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Sweden	SE	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Switzerland	CH	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Turkey	TR	x	x	x	x		x	x	x					x	x					x	x				
Ukraine	UA	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
United Kingdom	GB	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x		x	x	
United States	US	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Viet Nam	VN	x		x	x		x																		
Zambia	ZM	x	x	x	x		x																		
Zimbabwe	ZW	x	x	x	x	x	x	x	x		x	x		x	x		x	x		x	x	x	x	x	x

Note: For predicting marginal effects, we include the following additional 124 countries: *AE* (Emirates), *AF* (Afghanistan), *AG* (Antigua & Barbuda), *AL* (Albania), *AM* (Armenia), AO (Angola), *AZ* (Azerbaijan), *BA* (Bosnia & Herzegovina), *BB* (Barbados), BD (Bangladesh), BF (Burkina Faso), *BH* (Bahrain), BI (Burundi), BJ (Benin), *BN* (Brunei), BO (Bolivia), *BS* (Bahamas), *BT* (Bhutan), BW (Botswana), *BY* (Belarus), *BZ* (Belize), *CD* (Congo, Dem. Rep.), CF (Central African Republic), CG (Congo, Rep.), CI (Cote d'Ivoire), CM (Cameroon), CO (Colombia), CR (Costa Rica), *CV* (Cap Verde), CY (Cyprus), *DJ* (Djibouti), *DM* (Dominica), DO (Dominican Republic), DZ (Algeria), EC (Ecuador), ET (Ethiopia), FJ (Fiji), *FM* (Micronesia), GA (Gabon), GD (Grenada), GH (Ghana), *GM* (Gambia), *GN* (Guinea), *GQ* (Equatorial Guinea), GT (Guatemala), *GW* (Guinea-Bissau), GY (Guyana), HN (Honduras), HT (Haiti), IL (Israel), IN (India), IQ (Iraq), JM (Jamaica), *KG* (Kyrgyz Republic), *KH* (Cambodia), *KI* (Kiribati), *KM* (Comoros), *KN* (St. Kitts & Nevis), *KW* (Kuwait), *KZ* (Kazakhstan), *LA* (Lao), *LB* (Lebanon), *LC* (St. Lucia), LK (Sri Lanka), LR (Liberia), *LS* (Lesotho), *MD* (Moldova), *ME* (Montenegro), MG (Madagascar), *MH* (Marshall Islands), *MK* (Macedonia), ML (Mali), MM (Myanmar), *MN* (Mongolia), *MO* (Macao), MU (Mauritius), *MV* (Maldives), MZ (Mozambique), *NA* (Namibia), NE (Niger), NG (Nigeria), NP (Nepal), *NR* (Nauru), *OM* (Oman), PA (Panama), PE (Peru), PK (Pakistan), *PS* (West Bank & Gaza), *PW* (Palau), PY (Paraguay), *QA* (Qatar), *RS* (Serbia), RW (Rwanda), SA (Saudi Arabia), *SB* (Solomon Islands), *SC* (Seychelles), SD (Sudan), SL (Sierra Leone), SN (Senegal), SO (Somalia), *SR* (Suriname), *SS* (South Sudan), *ST* (Sao Tome and Principe), SV (El Salvador), SZ (Swaziland), TD (Chad), TG (Togo), TH (Thailand), *TJ* (Tajikistan), *TL* (Timor-Leste), *TM* (Turkmenistan), TN (Tunisia), *TO* (Tonga), TT (Trinidad and Tobago), *TV* (Tuvalu), TZ (Tanzania), UG (Uganda), UY (Uruguay), *UZ* (Uzbekistan), *VC* (St. Vincent and the Grenadines), *VU* (Vanuatu), *WS* (Samoa), *XK* (Kosovo), *YE* (Yemen). For ISO2 codes displayed in *italic* font, we do not observe the level of patent protection; this also includes *EE*, *GE*, *HR*, *LV*, and *SI* from above.

TABLE B2. Correlation matrix of inequality measures, population, and GDP

	$GINInet_{lt}$	$GINImar_{lt}$	$TOP10\%\text{-share}_{lt}$	$TOP5\%\text{-share}_{lt}$	$TOP1\%\text{-share}_{lt}$	$\log(POP_{lt})$	$\log(GDP_{lt})$
$GINImar_{lt}$	0.639***						
$TOP10\%\text{-share}_{lt}$	0.372***	0.289***					
$TOP5\%\text{-share}_{lt}$	0.480***	0.287***	0.974***				
$TOP1\%\text{-share}_{lt}$	0.588***	0.368***	0.860***	0.946***			
$\log(POP_{lt})$	0.375***	0.150***	0.167***	0.142**	0.102		
$\log(GDP_{lt})$	−0.106***	−0.020	0.218***	0.169**	0.088	0.373***	
$\log(GDPpc_{lt})$	−0.289***	−0.092***	0.132	0.085	0.029	−0.105***	0.884***

*** , **, and * indicates significance at 0.01, 0.05, and 0.1, respectively. Correlation matrix based on 414 observations on destination countries in 1980-2013. Bonferroni adjusted inference.

Table B3. Pair-wise correlations among family-forward citation measures

	log(PREVIOUS$_{it}$)	log(CURRENT$_{it}$)	log(AVSUB$_{it}$)	log(NODESTINS$_{it}$)
log(CURRENT$_{it}$)	0.931***			
log(AVSUB$_{it}$)	0.957***	0.911***		
log(NODESTINS$_{it}$)	0.688***	0.622***	0.655***	
log(NOSIMULTDESTINS$_{it}$)	0.584***	0.696***	0.577***	0.871***

*** indicates significance at 0.01 applying Bonferroni adjusted inference. Correlation matrix based on 4,506,447 observations on family-citations for 1,559,877 different DOCDB-families for innovations made in 1980-2013.

Table B4. Factor loadings on first principal component

Variable loadings on PATVAL$_{it}$	
log(PREVIOUS$_{it}$)	0.468***
	(0.000)
log(CURRENT$_{it}$)	0.468***
	(0.000)
log(AVSUB$_{it}$)	0.464***
	(0.000)
log(NODESTINS$_{it}$)	0.423***
	(0.000)
log(NOSIMULTDESTINS$_{it}$)	0.410***
	(0.000)

*** indicates significance at 0.01.

Figure B1. Densities of imputed (solid) versus original data (dashed)