Do privacy laws affect the location decisions of internet firms? evidence for privacy havens

ROCHELANDET, Fabrice and TAI, Silvio H.T.

Department of Economics, University of Paris Sud

2012

Do Privacy Laws Affect the Location Decisions of Internet Firms? Evidence for Privacy Havens

Fabrice Rochelandet[○] & Silvio H.T. Tai[°]

[○] ADIS, Université Paris-Sud, France

[°] PPGE/PUCRS, Brazil & ADIS, Université Paris-Sud, France

Abstract

This paper empirically studies the location decisions of internet firms when they face high legal standards of privacy protection. Many factors might influence them: technological spillovers, lower taxation, and so on. Internet firms can also arbitrate national differences and many of them actually locate their activity in order to escape from national laws they consider over-stringent. In the current stage of development of the internet – the so-called Web 2.0 – the ease of access to personal data proved to be strategic input. So the more a jurisdiction makes collecting and using these data easy, the more attractive the country is, if all other things remain constant. One way for a firm to avoid such legal restrictions is to locate or to expand its business in less privacy protective countries. Our empirical results support this 'no-privacy haven' hypothesis. In particular, we highlight a new privacy paradox according to which the more stringent certain online privacy laws are, the more they induce firms to locate their business in less stringent countries, and finally the weaker actual privacy protection on the internet is.

JEL codes: K29, F23, L86

Key words: Privacy, Firms' Location, Internet

We would like to acknowledge the helpful comments of our colleague Jean-Michel Etienne, the participants in the ADIS research seminar, and two anonymous referees. This work was supported by the French National Agency of Research, research program ANR-CONTINT-ESPRI. The authors gratefully acknowledge Google France for partially supporting this research.

1. Introduction

Privacy regulation is one of the most important debates in internet policy. For instance, Goldfarb & Tucker (2011) suggest that the EU's e-Privacy Directive affects the performances of online advertising but more generally internet-based activities because of the importance of advertising revenues in supporting the development of the digital economy. A delicate compromise is to be found concerning legal restrictions on the commercial use of personal data. On the one hand, many online businesses are suspected of invading the privacy of individuals causing them various forms of harm. On the other hand, over-stringent rules can impede innovation on the web – targeted marketing, online customization – and perhaps impact internet firms' location decisions in favor of less protective countries. In turn, this could entail a race to the bottom leading to laxer privacy to the detriment of internet users. This may be a new privacy paradox, i.e. stronger local privacy laws entail a reduction in global privacy protection for internet users. This new paradox can be explained by the contrast between national jurisdictions and the 'global village' nature of the internet.

In this paper, we pay particular attention to this dimension. We focus on privacy laws and ask how they affect the location decisions of internet firms in an international setting. In particular, we focus on the Web 2.0 services as the biggest users of personal data. Thus, our

main question is whether more stringent privacy laws in one country can lead firms to locate or to expand their business into less protective countries.

One naïve approach might consist in considering the digital economy as a space where geography does not matter (Choi & Suh, 2005). As a consequence, no significant physical or technological barriers would prevent online activities from expanding into digital networks while remaining located in their respective country of origin. However, the facts suggest that internet firms tend to locate their business in countries that are often different from their national origin. Many factors may influence their location decisions: high-skilled and cheaper workers, spillovers and agglomeration effects, more favorable taxation, innovation policies, cluster development programs, and so on. Differences in national jurisdictions can also play an important role. Internet firms are often purported to locate their business overseas in order to exempt themselves from the jurisdiction of local law and thus benefit from a 'legal haven'. For instance, illegal file-sharing networks often locate their servers in copyright-free countries ('copyright havens' or 'copyright hells'). In the field of privacy, Facebook located its European headquarters in countries like Ireland but not in France, where legal protection of personal data is more stringent.

This paper shows that privacy and the legal protection of personal data explain the location decisions of internet firms. This result suggests a 'no-privacy haven' effect that induces them to locate their business in more privacy lax countries.

The rest of this paper is organized into five sections. The next section provides a brief survey of the literature that seeks to explain firms' location decisions. The third section presents our hypotheses. Section four describes the method: variables used and the econometric model. Section five presents the main results. Section six envisages some policy implications.

2. Literature Review

The internet has been thought of as an increasingly smaller place for business by virtually abolishing transportation costs (Cairncross, 1997) - services are delivered online – and factor-mobility is enhanced thanks to telecommuting and videoconferencing, for instance. For internet users, national boundaries are transparent: Users can surf from one website to another, whatever the location of servers, websites and internet firms and, most of the time, without being informed of their nationality. As is the case with firms, managers, trade partners, and high-skilled workers can communicate and travel across national borders in a low-cost, quick and easy fashion. This, in turn, has abolished government barriers for many countries and thus facilitates the operation of online businesses.

However, as has already been mentioned, many internet firms have multiple locations at the international level. This suggests that ITs have mixed effects on the cost of distance (Venables, 2001). ITs eliminate transport costs for online services – such as music or airline ticketing – that can be codified and then no longer need to be close to consumers. But the elimination of this kind of constraint does not mean that all other constraints are eliminated. Internet firms will relocate parts of their business according to lower costs and other non-pecuniary advantages. The location decisions of firms can be explained by many factors prevailing at the international and regional levels. This can be conceived as a (simple) sequential process: the firm first chooses to locate or to expand its business in a specific country/international area and then it selects a specific location inside the country/area (Mayer & Muchielli, 1999). Alternatively, the firm can select directly among different regions and cities according to their respective attractiveness. But *whatever the means of choosing a location*, national determinants should always prevail and influence the actual decision of the

firm by affecting its costs and profits[1]. The legal protection of personal data can actually impact on these variables.

Privacy economics: the impact of legal standards on firms' revenue

Since the 1970s many economic studies have investigated the effects of privacy on social welfare. In particular, economists have discussed the impact of this legal protection on market efficiency. For some scholars of the Chicago School, it reduces the quantity of available information on consumers, workers, job seekers, debtors, and so on. Thus, social welfare might be affected because of an inefficient matching of supply and demand (Posner, 1980, Stigler, 1978). However, with the advent of social networking services this perspective proves irrelevant since individuals tend to divulge an increasing volume of private information about themselves.

More recently, scholars have identified different categories of costs incurred by firms when they have to comply with privacy legislation[2]. Acquisti (2010) distinguishes the costs of protecting data from those associated with disclosing them. In both cases, a higher protection of personal data might lead to a reduction in net gains of the firms that collect and use them. More protection can result in a decrease of the revenues of firms by preventing them from learning more about individuals and thus extracting more surplus from consumers, better personalizing their services, minimizing their advertising costs by better targeting customers, decreasing their inventory risks, better selection of job seekers, and so on. In addition, privacy laws can restrict the ability to resell consumer data, in particular with firms located in other countries. In some cases, such as that of France, specific administrative authorities control the international transfer of personal data. Unequal access to personal data can also prove a competitive disadvantage for firms that cannot collect them or that face restrictions.

A significant increase in privacy protection can also lead to higher direct costs due to the expense of legal compliance. Breaking the privacy law can result in firms paying fines, legal expenses and damages. And being sentenced for a breach of privacy might constitute another opportunity cost in terms of a negative reputation leading to a reduction in sales or stock market value. To avoid such costs and their liabilities, firms might have to invest resources by hiring lawyers, adapting their contracts to the law, designing privacy charts to comply with the law and securing their databases. At an international level, multinational firms have to ensure that their terms and conditions conform to local privacy rules. All these costs are positively correlated with the legal standards of privacy and are different according to the regulation (Swire, 1997; Tang et al., 2008).

From the sole viewpoint of firms, (more stringent) privacy rules are likely to decrease their profits when collecting and using personal data. Therefore, when facing different levels of privacy protection, one can expect them to lobby or arbitrate in favor of the less stringent ones. At the international level, this can turn out to be a decision in terms of location in different countries. Being profit maximizers, they will locate to countries/areas where the costs of production are minimized. In particular, internet firms are presumed to be the most

[1] Simply stated, firms will be assumed to systematically avoid both heavy taxation and wars, whatever talent and knowledge spillovers might prevail there.

[2] Of course, firms can also benefit from the legal protection of personal data (Hahn et al., 2001). For instance, direct marketers and online merchants can improve their business processes when those rules aim at fighting against spamming or customer identity theft. Moreover, collecting too much additional information can be costly for firms and market efficiency (Hermalin & Katz, 2006). Conversely, lax protection can induce mistrust from individuals that might consider firm strategies as too intrusive.

sensitive to the privacy laws because they use large amounts of personal data to run their online business.

'Classical' factors for firms' location decisions

Traditionally, firm location decision refers to foreign direct investment literature analyzing the impact of globalization. The progressive removal of trade barriers enables firms to develop international activities and strategies. In particular, factors such as wage and productivity levels, size of foreign markets, distance from the home country and income taxation influence their decision to produce abroad rather than producing at home for export.

A foreign direct investment is conventionally classified according to the strategy adopted by the firm. Vertical integration exploits favorable conditions of factor prices in the country invested in whilst horizontal integration prevents trade costs from serving the host market locally. A single firm may however follow a complex strategy investing in both vertical and horizontal modalities. To illustrate this, American subsidiaries in Canada are composed of 12% of horizontal FDI, 19% of vertical FDI and 69% of complex integration strategy of investment (Feinberg and Keane, 2003). Yeaple (2003) shows that vertical and horizontal FDI present a complementary relationship and, for intermediate levels of transport costs, only complex integration is worthy because of this relationship. Vertical integration allows for a lower cost per unit which expands sales, raising the firm´s capacity to engage horizontal FDI. Grossman, Helpman and Szeidl (2006) extend the analysis of Yeaple (2003) applying the heterogeneity of firms developed by Melitz (2003). They find several configurations of locations, according to the fixed costs incurred offshore. A general result is that the most productive firms tend to locate in countries with lower factor prices whilst the least productive firms stay in the home country.

One can think of the web 2.0 sector as institutionally dependent. The products of internet firms require a low level of legal restriction that is ultimately determined by the institutions of a given country. Levchenko (2006) shows that countries with favorable institutions for this kind of product present a comparative advantage. Consequently, countries with low restrictiveness on the exploitation of data present an expansion of the internet firm sector, accommodating global demand[3].

Spillover effects and location decisions of high-tech firms and R&D activities

Recent theoretical research into new economic geography studies the determinants of production patterns and firms' location decisions in an international setting (Krugman, 1991, Fujita, Krugman & Venables, 2001, Fujita and Thisse, 2002). Pecuniary and technological externalities influence firm location decisions. Pecuniary externalities refer to interactions generated by market and pricing mechanisms. Technological externalities are associated with factors such as the face-to-face exchange of information, access to intermediate inputs, high-skilled workers and specialized technology suppliers, as well as market size. In particular, firms capture those knowledge spillovers through their proximity to other organizations. The notion of 'cluster' (Porter, 1998) encompasses these aspects.

In particular, many factors affecting foreign R&D location decisions in high-tech sectors and R&D activities have been identified (Jones & Davis, 2000). Multinational firms locate and expand their R&D activities because they need to acquire and leverage innovative capacity

[3] Levchenko (2006) considers the quality of the contracting institution, but the author points to other institutions such as the rigidity of labor market and union power (p. 797).

abroad while benefiting from economic spillovers generated by the co-location of competitors and business partners. The ability to access local scientific talent, creativity and technology, as well as the strategic possibility to observe competitors' R&D activities are key determinants. Demand factors also play an important role, i.e. being close to final markets and more responsive to local opportunities.

Symmetrically, governments and local authorities seek to attract innovative firms and R&D activities in their respective countries and regions. Thus, the quality of the political environment is crucial. The level of appropriability regimes is not only determined by the quality of property law - in particular, patents and copyrights – but also by the level corruption and political risks. All these factors influence a firm's decision about where and how much to expand their R&D activities. In some cases, governments can ameliorate the institutional environment of their countries by developing specific aspects such as scientific capability and telecommunications infrastructure and services. In so doing, governments can enter into a race to the top or… to the bottom to attract more R&D thus fostering national growth.

'Pollution haven' hypothesis applied to privacy regulation: A race to the bottom?

For certain sectors, the differential in terms of national regulations constitutes a factor that can explain the location and expansion of businesses. Firms will therefore engage in regulatory arbitrage, moving to countries with the lowest legal standards. Symmetrically, less stringent laws and regulations ranging from labor laws to bioethics can be considered a comparative advantage for countries that adopt them. States seeking to attract or fearing to lose firms can choose to soften regulatory standards. In particular, this 'race-to-the-bottom' hypothesis has been analyzed in the field of labor and environment regulation but without appreciable evidence of a strong correlation between legal standards and location decision (Drezner, 2006). However, these studies of the race-to-the-bottom hypothesis are based on a macroeconomic perspective and few investigations focus on the sector level.

The 'pollution haven' hypothesis stipulates that national jurisdictions can turn into pollution havens since dirty industries can relocate in response to differences in regulatory stringency (Brunnermeier & Levinson, 2004). In the same way, by assimilating personal data users as (potential) privacy polluters[4], privacy regulations can influence their location decision. But in the case of privacy legislation, the 'no-privacy haven' hypothesis is slightly different from the 'pollution haven' framework. Differences among countries could be initially thought of as being unintentional from the perspective of lawmakers and regulators. They could result from different historical processes or culture or political contexts. And on this basis, firms can ground their location decisions by taking into account differences in the laxity of privacy national laws.

We can also infer from all these aspects that features of the internet firm and its online activity might influence its location decisions. The size of the firm is an important determinant, even though it does not act so much as an incentive but more as a capability factor. The nature of the business model can impact on the need for personal data. In other words, the more intrusive the business model is, the more appealing will be a regulation based on a weaker protection of online privacy. This point refers to the above-mentioned 'no-privacy haven' hypothesis.

[4] The analogy we make here between natural environment and internet seems all the more relevant since those two 'milieu' make the production functions of firms, the utility functions of individuals, and the countries' welfare interdependent, and thus they are both likely to create negative externalities.

3. Hypotheses

According to the literature reviewed in the previous section, the main factors that may explain the firm location decision can be grouped into two main categories.

The first category evaluates the factors associated with the attractiveness of the country. The impact of this first category is captured through four kinds of variables: supply factors (the size of the high-tech sector; the size of the internet infrastructure), demand (population penetration of internet use), competitive/herding factors (the location of other internet firms, i.e. the number of internet firms located in the same country), and the institutional environment (the quality of the institutional environment; the stringency of privacy protection).

We could expect a positive relationship between both an access to a much more developed telecom infrastructure and the decision to locate in that country. A more extensive internet connectivity is presumed to benefit internet firms.

Hypothesis H1.1: *Internet firms' decisions to locate depend positively on the size of the telecom infrastructure.*

Internet firms are presumed to choose on the basis of the differences in taxation between locations since a higher average rate will impact negatively on their profits. However, in the case of internet firms, this hypothesis can be challenged because most of the profits of the firm's profits can be made in other countries. For instance, a web audience can be created in one country and advertising revenue paid in another country.

Hypothesis H1.2: *Location choices depend negatively on tax* ***levels.***

Internet firms are a specific category of high-tech business. They need access to a highly skilled labor market and proximity to other firms in order to observe them.

Hypothesis H1.3: *The size of the high-tech sector in the host country and spillover effects are positively associated with the internet firm's decision to locate or to expand its business in that country.*

As already noted, fiscal or labor advantages are not always sufficient grounds upon which to base a location decision. Firms have to consider institutional stability too. Any location decision implies a certain level of investment that can represent a substantial loss if firms have to leave a country. Thus, firms have to take into account the risks inherent in politically unstable countries. At first glance, due to the immaterial nature of their business, internet firms are characterized by a much higher degree of footlessness than other industries. Actually, it depends upon the precise nature of the investment: creating a marketing department or implementing servers and communication networks does not generate the same level of investment. In any case, we presume that institutional stability impacts on the location decision of internet firms.

Hypothesis H1.4: *Evidence of institutional quality is positively associated with firm location decision in that country.*

According the literature mentioned in the previous section, internet firms are presumed to be sensitive to the level of legal protection of personal data. In particular, such rules can hamper not only the methods of collecting and processing those information goods, but can also impact innovation in online services offered by internet firms and hence both their profits and

costs. Nonetheless, from our point of view, the 'death-of-distance' thesis might be true concerning the collection of personal data: whatever the restriction laws impose upon the method used to gather data, individuals are free to divulge what they want about themselves. So privacy rules about collection might not impact on the location decisions of internet firms. However, this is not the case with the legal constraints on the use of data collected and on the ways of building online services using this data. These privacy rules can hamper innovation and the way of doing business. They might explain the location decisions of internet firms. In turn, the more useful and innovative individuals judge online services, the more information they are willing to divulge about themselves. So privacy rules that regulate the processing and use of personal data prove crucial for firms when they decide to locate their business in various countries.

Hypothesis H1.5: *Evidence of privacy protection (high stringency) is negatively associated with firm location decision in that country.*

The second category of variables refers to features of the firm itself. All firms perform online business, but their activities can be more or less intensive in personal data. Thus, firms are expected to be sensitive to privacy rules in proportion to the intrusiveness of their online business.

Hypothèse H2: *The greater the privacy intrusiveness of online business, the more likely the internet firm is to decide to locate or to expand its business in countries with weaker privacy protection.*

4. Methodology

4.1 Econometric specification

We investigate choice-specific determinants of industrial location through a Poisson estimator. In opposition to discrete choice models like the multinomial logit, the Poisson provides a more tractable model. Guimarães et al. (2003) show that both the conditional logit and the Poisson can present the same log likelihood and coefficients.

Although the traditional estimator used for discrete multiple choices is the multinomial logit, we prefer to apply a Poisson estimator. With the Poisson we estimate the quantity of firms located in each destination with country specific dependent variables. Below we sketch the econometric model proving the equivalence between these two estimators.

The utility of a firm k locating to country l is:

$$\pi_{kl} = \boldsymbol{\beta}' \mathbf{z}_{kl} + \varepsilon_{kl} \qquad (1)$$

Where β is a vector of unknown parameters, z_{kl} is a vector of explanatory variables, and ε_{ij} **is** a random term.

Adopting the appropriated distribution of errors, the probability that a given firm will choose a region l is expressed by a multinomial logit function (McFadden, 1974). Each location is identified by l and the total number of locations is equal to L. Thus, the probability of a firm k choosing location l is:

$$p_{kl} = \frac{e^{\boldsymbol{\beta}'\mathbf{z}_{kl}}}{\sum_{l}^{L} e^{\boldsymbol{\beta}'\mathbf{z}_{kl}}} \tag{2}$$

The log likelihood for the conditional logit is expressed by:

$$L_{cl} = \sum_{k=1}^{K}\sum_{l=1}^{L} d_{kl}\log(p_{kl}) \tag{3}$$

where d_{kl} is equal to one if firm i chooses location l and is otherwise zero.

While multinomial logit functions are a conventional way to estimate location choices, we follow Guimarães et al. (2003) with the application of a Poisson estimator. This procedure provides a more tractable model, avoiding the problem of non-linearity.

It is interesting to consider at this point that firms are aggregated by sector, which is the way the empirical analysis will proceed. Then, considering the quantity of firms q_{sl} from sector s locating in country l to be Poisson distributed:

$$E(q_{ls}) = e^{(\boldsymbol{\alpha}'\mathbf{d}_{ls} + \boldsymbol{\beta}'\mathbf{z}_{ls})} \tag{4}$$

where [α, β] are parameters to be estimated and d_{ls} is a vector of s dummy variables for each sector s.

The Poisson model's likelihood is expressed as:

$$L_p = \sum_{s=1}^{S}\sum_{l=1}^{L}\left[-\exp(\boldsymbol{\alpha}'\mathbf{d}_{ls} + \boldsymbol{\beta}'\mathbf{z}_{ls}) + q_{ls}(\boldsymbol{\alpha}'\mathbf{d}_{ls} + \boldsymbol{\beta}'\mathbf{z}_{ls}) - \log q_{ls}!\right] \tag{5}$$

Considering the first order conditions with respect to α_s:

$$\frac{\partial L_p}{\partial \alpha_s} = \sum_{l=1}^{L}\left(q_{ls} - e^{\alpha_s + \boldsymbol{\beta}'\mathbf{z}_{ls}}\right) = 0 \quad \Rightarrow \quad \alpha_s = \log\left(\frac{q_s}{\sum_{l=1}^{L} e^{\boldsymbol{\beta}'\mathbf{z}_{ls}}}\right) \tag{6}$$

where $q_s = \sum_{l=1}^{L} q_{ls}$

Linking equation (6) to equation (5), the Poisson likelihood becomes:

$$L_p = \sum_{s=1}^{S}\sum_{l=1}^{L} q_{ls}\log(p_{ls}) - K + \sum_{s=1}^{S} q_s \log q_s - \sum_{s=1}^{S}\sum_{l=1}^{L}\log n_{ls}! \tag{7}$$

The last two terms of equation (7) are constants. The first term is the log-likelihood of the conditional logit. Estimated coefficients are the same in both models. Estimated regressions are carried out using the Poisson model with quantities as the dependent variable.

A very restrictive characteristic of the Poisson distribution is the equality between the mean and variance. However, if data violates this condition, it does not definitely invalidate the results. Firstly, in such cases, coefficients are nonetheless consistently estimated. Secondly, the quasi-maximum likelihood estimation (QMLE) can adjust standard errors (see Wooldridge 2001, section 19.2).[5]

4.2 Sample and variables

Dependent variable

Our dependent variable is the Location Decisions of internet firms. Data on firms' location were collected considering the 124 most viewed firms on the internet according to the Alexa ranking. Alexa provides a measurement of the audience of internet sites by computing traffic indicators based on a three month moving average of aggregated historical traffic data from millions of Alexa toolbar voluntary users. The main indicator developed and published by Alexa is "Traffic Rank" which yields the position of the site in interest with respect to all the sites on the web.

We then produced original data on the locations of the corresponding internet firms. To construct the dependent variable, we collected the locations of the headquarters of 124 internet firms spread over a total of 76 countries. Obviously, additional information about the precise nature of the locations – servers, R&D units, marketing subsidiaries and so on – would improve our results concerning the calculations of firms when trying to avoid over-stringent privacy laws. It might help us to dissociate more accurately the relative effects of the tax system, spillovers… and privacy. However, because of a lack of necessary data that are often of a strategic nature, it proves impossible to formally test for more specific locations. Furthermore, as such, the location of the headquarters can be considered as relevant to an explanation of firms' choices when trying to avoid over-stringent privacy laws.

The locations of internet firms are aggregated according to 10 sectors ranging from 'search engines & web portals', 'online networking services', 'web hosting & sharing services', 'online press & information services', 'infomediaries', 'e-commerce', 'corporate websites', 'online administration', 'e-advertising' to 'online games'. Therefore, the dependent variable is the quantity of firms of a given sector *s* located in a given country *l*. These data were matched with the privacy index derived from the *Privacy International Survey*. Altogether, we obtain 152 observations in 37 countries, i.e. 37 'privacy-rated' countries times 10 activities, less the number of 'holes' in terms of lack of localization.

Explanatory variables

Privacy index

Data on privacy were collected from the *The Privacy & Human Rights Report* published by Privacy International (2007). The privacy index surveys the state of surveillance and privacy protection in 70 countries. It constitutes the most comprehensive report published on this topic. It is built by an international network of experts ranging from researchers, human

[5] Assuming that the variance is proportional to the mean, standard errors are corrected multiplying by the sum of Pearson residuals divided by the degree of freedom.

rights advocates and journalists to regulators and policy-makers around the world[6]. In particular, most of the criteria used to build this score are likely to affect the location decisions of internet firms: for instance, 'Privacy enforcement', 'Data sharing provisions', 'Constitutional protections', and 'Statutory protection ' (see Appendix A1 for details).

'Classical' variables: economy size, labor costs and profit taxes level

Based upon the existing literature on determinants of firms' location and FDI, our regressions also include proxies for taxes on business profits.[7] These data were collected from World Bank Development indicators.[8]

Spillover effect and the internet

A variable evaluates technological externalities: a 'Business Technological Environment' index (source: Centre d'Etudes Prospectives et d'Informations Internationales[9]). In addition, we include the number of internet users that constitutes our proxy for measuring the general quality of the local telecom infrastructures. These aspects are potentially significant for internet firms looking for efficient communication networks. Furthermore, we test for the number of servers that is likely to influence the location decision because of better access to the internet.

Institutional quality

In our study we have two complementary proxies measuring institutional quality: 'Political Stability' and 'Transparency of public actions in the economic field' that were derived from the Centre d'Etudes Prospectives et d'Informations Internationales (Cepii). The first variable measures the durability of a political regime and thus, the quality of the commitments of a given government. It might be sensitive for online press businesses and web hosting content that can prove political. The second variable is essential for an accurate evaluation of the economic environment, on the one hand, and the anticipation of the long-term strategy of a government in terms of economic policies concerning taxes, budget, money supply, exchange rate, and so forth, on the other. The transparency of economic policy assures accessibility and liability to information, improving firms' perception of institutional quality.

Table 1 reports descriptive statistics for the independent and dependent variables.

[6] "The ranking assesses the key areas of surveillance and control, and will identify mechanisms of protection that have failed to operate according to the letter and spirit of the national and international privacy protections. It will concentrate on policy development issues, inadequacies in the consultation process, legal protections (or lack of them), the impact of surveillance on democratic institutions, changes to the nature of society and the implications for individual freedoms and autonomy." (Privacy International, 2007)

[7] The size of the economy measured by the GDP is correlated to the quantity of servers at 95%. This paper uses the latter variable as it constitutes a key element regarding telecom infrastructure required by internet firms.

[8] Although the cost of labor is an important control, available statistics do not detail industries compatible with the internet ones, providing data on sectors like "Manufacture of textiles" (see http://laborsta.ilo.org/default_F.html).

[9] See Appendix A1 and the Cepii website (www.cepii.fr) for more details about the construction of this variable.

Table 1: Descriptive statistics

Variable	Nb of obs.	Mean	Std. Dev.	Min	Max
Firms' location	152	2.59	4.52	1	38
Privacy index	152	2.17	0.51	1.3	3.1
Total tax rates (% commercial profits)	152	49.37	15.13	22.6	108.1
Political stability	152	3.92	0.26	2.57	4
Transparency of public economic actions	152	3.56	0.56	1.43	4
Business technological environment	152	3.22	0.68	2.27	4
Internet users (per 100 people)	152	0.62	0.23	0.053	0.918
Servers	152	3.74 e+07	1.07 e+08	188000	4.39 e+08

5. Results

Table 2 reports the results for the determinants of the location of firms obtained using a Poisson QMLE estimator.[10] The dependent variable is the quantity of firms in a given sector located in each country. The coefficient for the level of privacy regulation is statistically significant at the 1% level for all regressions. Column (1) shows the result only for the privacy variable (with sector fixed-effects results omitted). As expected, the coefficient is negative and significant. The higher the privacy regulation in a given country, the lower the number of web firms locating in this country. To have an idea, a reduction of one standard error of privacy variables explains the location of one firm.[11]

This first regression does not involve any control variables and the magnitude of the coefficient of privacy may be biased. The following regressions progressively add control variables and the magnitude of the coefficient of privacy varies. Column (2) adds the level of profit taxes (from commercial operations) in the regression. This variable does not present a statistically significant coefficient. It may be explained by the fact that web companies can have operations and financial transactions in different countries.

An important variable of control is the political stability of a country. As Tole and Koop (2008) point out, from the firm's point of view there is a trade off between the level of regulation and the stability a given country can offer. For example, some less developed countries may offer lower regulation but they cannot necessarily assure stability for businesses in the long run. Regression (3) introduces this variable that is positive and significant. Under this control, the magnitude of the coefficient of privacy rises to -1.43.

Column (4) adds other controls like the Transparency of Public Actions in the economic field and the Business Technological Environment. Both present the expected results: firms locate where economic actions are transparent and where the business and technological environments are important. The former is not statistically significant while the latter is highly statistically significant and robust (except in the last regression).

Regression (5) introduces the number of internet users for and the logarithm of the number of servers hosted in each country. The internet users´ coefficient does not present a significant coefficient. One possible explanation is that users frequently set up webpages in different

[10] As a robustness test, table A4 provides the result of the same regressions using a negative binomial estimator. Results are very similar.

[11] This number is obtained considering the average level of the location variable (2,59) and the standard error of the privacy variable (0,51). As the coefficient of column 1 is equal to 0.97, 1,28=0.97*0,51*2,59

locations. If the quantity of internet users does not matter for the location of firms, the quantity of servers seems fundamentally connected to this decision. The thinking behind for these results is also convergent; users can access the internet from anywhere but servers are location-specific. The increase of 10% in the number of servers in a country is related to a 3,8% rise in internet firms, although causality cannot be determined.

Finally, regression (6) replicates regression (5) discarding the United States from the sample. This test is very important, given that 25% of firms locate to this country. The United States also presents a low level of privacy restriction and one could imagine that another factor could explain simultaneously web firms' location and the degree of privacy in the USA. Indeed, the coefficient of privacy is reduced in regression (6) but is still statistically significant at the 1% level. If one considers the United States, it raises the correlation between firms' location and privacy, but this also applies to the data from the other countries.[12]

Table 2 – Econometric Results – Poisson QMLE

Dependent variable: Firms' Location	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	All	All	All	All	No USA
Privacy	-0.97***	-0.97***	-1.43***	-1.58***	-0.67***	-0.36***
	(0.19)	(0.19)	(0.19)	(0.15)	(0.12)	(0.11)
Ln(Tax)		0	0.26	0.87***	-0.45**	-0.09
		(0.33)	(0.29)	(0.25)	(0.22)	(0.19)
Political Stability			2.21***	1.53***	0.67**	0.49**
			(0.54)	(0.46)	(0.27)	(0.23)
Transparency of Public Economic Actions				0.04	0.25**	0.1
				(0.15)	(0.12)	(0.10)
Business Technological Environment				0.80***	0.22**	0.06
				(0.10)	(0.10)	(0.09)
Internet users					-0.47	0
					(0.30)	(0.27)
Ln(Servers)					0.38***	0.13***
					(0.03)	(0.05)
Constant	3.26***	3.28**	-5.49**	-7.65***	-5.64***	-2.52**
	(0.42)	(1.29)	(2.39)	(1.77)	(1.13)	(1.09)
Observations	152	152	152	152	152	142

Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. As data presents overdispersion, standard errors are corrected by the Pearson chi-squared statistic divided by the degrees of freedom. All regressions include sector fixed-effects.

Table 2 presents very significant and robust results for the impact of privacy on the location of internet firms. The coefficient of privacy consistently produces a confidence level of less

[12] Table A5 in the Appendix tests the web 2 firms´ specificity including a dummy for these sectors, however its coefficient is not statistically significant. Table A5 also provides a robustness test including the share of firms per country minus the share of firms in the sector per country (in the sample). These results should be taken into account carefully because of the endogeneity these regressions present. Nonetheless, the privacy index variable remains statistically significant in these tests.

than 1%. Nonetheless, controls are not exhaustive and reverse causality may be an issue.[13] One can assume that multinationals, once located in a given country, may look for lower levels of privacy regulation through lobbying pressures. However, the privacy variable is based on macro measures[14] and may contain errors. [15]

Missing variables, reverse causality and measure errors are a source of endogeneity that can be addressed by the use of instrumental variables. We apply two variables that are correlated to the privacy regulation level and are not correlated to the firm´s location. The first one is the quantity of civil society organizations in a given country. The United Nations Organization (UNO) attributes a consultative status for non-governmental organizations that "gives access to intergovernmental processes at the UNO dealing with economic and social development, gender issues, sustainable development, small arms, and human rights". [16] The quantity of civil societies in a country may have a strong negative correlation with the level of privacy regulation; countries that have more tolerant communities regarding non-governmental organizations are more likely to have lower levels of privacy regulation. Indeed, the regression of the Privacy variable on the Civil Society variable gives a t-statistic of -4.99 and an r^2 of 0.1425. However, there is no reason for a firm to take into account the quantity of civil society organizations in a country for the choice of location.

The second instrumental variable is the Freedom of the Press. We also expect that in countries where the press is freer, the privacy regulation level is lower. The regression of the privacy variable on the Freedom of the Press variable reports a t statistics of -7.26 for press freedom and an r^2 of 0.2599. While the first instrument seems to fulfil the exclusion hypothesis, the second one sometimes fails in this task. One of the business models we have in our data is the fourth one: "online press & information services". The Freedom of the Press variable could possibly impact upon the location of online newspapers. We address this issue dropping observations from this business model of the dataset. Table 3 reports results for the sample without business model four. We also provide in table A6 in the Appendix the results for the whole sample.

Table 3 reproduces regressions (5) and (6) of table 2 using instrumental variables. Columns (1) and (2) apply the Civil Society Variable as instrument, columns (3) and (4) apply the Freedom of the Press variable instead. Finally, columns (5) and (6) apply both the Civil Society and Freedom of the Press variables as instruments. Both the coefficient of privacy and its standard error are higher than OLS estimations. Privacy does have a negative impact on firms´ location even treating potential endogeneity. Moreover, higher coefficients indicate a measurement error in privacy data. Conditional on the exogeneity of the instruments, Wu-Hauman tests confirm endogeneity in OLS estimations and Sargan/Hansen tests confirm the relative exogeneity of the instruments.

Compared to the OLS interpretation of coefficients, in regression (5) a reduction of one standard error of the Privacy variable explains the location of two firms (instead of one firm,

[13] Alternative regressions include GDP and GDP per capita. Results are similar and can be provided under request.

[14] See for details: https://www.privacyinternational.org/article/leading-surveillance-societies-eu-and-world-2007.

[15] While these issues can be addressed by IV techniques, another problem remains. As we selected the 124 most viewed firms on the internet for this analysis, a selection bias can be associated with these data by construction.

[16] Data are available on http://esango.un.org/civilsociety/displayAdvancedSearch.do?method=search&sessionCheck=false.

in the corresponding OLS estimation). It is also interesting to note other changes in variables. The coefficient of Taxes becomes non statistically significant, while Political Stability, Business Technological Environment gives coefficients with higher magnitudes and significance levels.

Table 3 – Poisson QMLE Intrumental Variables

Dependent variable: Firms' Location	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	No USA	All	No USA	All	No USA
IV	Civil soc.	Civil soc.	Press freedom	Press freedom	Both	Both
Privacy	-1.17***	-0.75***	-1.02***	-0.63***	-1.09***	-0.75***
	(0.18)	(0.19)	(0.19)	(0.16)	(0.14)	(0.15)
Tax	-0.04	0.03	-0.41*	-0.05	-0.16	0
	(0.26)	(0.21)	(0.23)	(0.21)	(0.23)	(0.21)
Political Stability	1.10***	0.78***	0.94***	0.71***	1.06***	0.79***
	(0.20)	(0.18)	(0.21)	(0.18)	(0.18)	(0.17)
Transparency of Public Economic Actions	0.20*	0.13	0.30***	0.13	0.22**	0.14
	(0.11)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)
Business Technological Environment	0.25**	0.15*	0.30***	0.13	0.28***	0.16*
	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(0.08)
Internet Users	-0.37	-0.14	-0.63**	-0.13	-0.43*	-0.17
	(0.25)	(0.21)	(0.26)	(0.22)	(0.24)	(0.21)
Ln(Servers)	0.27***	0.16***	0.39***	0.16***	0.30***	0.16***
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)
Constant	-6.04***	-3.89***	-6.65***	-3.53***	-6.21***	-3.98***
	(0.97)	(1.09)	(1.04)	(1.06)	(0.93)	(1.04)
Observations	143	134	143	134	143	134
Wu-Hausman (p-value)	0.0000	0.0115	0.0037	0.0021	0.0000	0.0000
Sargan/Hansen (p-value)					0.8803	0.8798

Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. As data presents overdispersion, standard errors are corrected by the Pearson chi-squared statistic divided by the degrees of freedom. All regressions include sector fixed-effects. See first step in table A7 in the Appendix.

6. Discussion & Policy Recommendations

In this paper, we investigate how privacy regulation influences the location decisions of internet firms. Our empirical results suggest that, in addition to classical factors such as political stability and spillovers, firms arbitrate according to differences in terms of privacy protection. Interestingly, this result is not more prevalent for internet firms that offer Web 2.0 services, although their business models are mostly based on the exploitation of personal data. Moreover, our study gives empirical evidence to highlight a new 'privacy paradox'[17] according to which the more stringent certain privacy laws are, the more firms are induced to

[17] The first 'privacy paradox' refers to the gap between people's privacy attitude and their actual behavior when they divulge their personal information when making a transaction or using social networking services (see Acquisti, 2004).

locate their business in less stringent countries, and finally the weaker actual privacy protection on the internet is. To some extent, this process would also be a favorable condition for the regulation of competition.

Two legal implications can be formulated from our results.

(1) The effectiveness of privacy regulation on the internet –as a 'global' network of networks– could justify a significant effort on the part of states to harmonize their privacy laws. Such an international agreement could be settled according to well-specified minimum standards as with the Berne Convention for the Protection of Literary and Artistic Works and the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) in the field of intellectual property. However, unlike these agreements, it seems that the governments that are the most likely to enable the prisoner dilemma underlying the 'new privacy paradox' to be overcome are also those which fully benefit from the locations of internet firms in terms of comparative advantages. So individuals might suffer for a long time from this race to the bottom. In addition, legal harmonization can have two negative effects: first, internet firms would be deprived of an opportunity to innovate by arbitrating between different legal orders; second, the harmonized standards could be fixed according to the "relative weights of interest groups asking for legal standards that benefit themselves, which can create inefficiencies if those standard rules prove to be unbalanced and generate losses of welfare due to legal irreversibilities over time" (Ribstein & Kobayashi (2006). In the case of privacy, this risk seems to be prevalent because this social norm is currently moving and gqsit gives rise to many counter-claims from civil society, telemarketers, internet firms, and so on.

(2) The impact of privacy reinforcement and/or harmonization in terms of social welfare seems to be very tricky to evaluate. On the one hand, governments look for comparative advantages. Privacy protection level might be one of them if softening it can lead to greater benefits from more innovative services. On the other hand, individuals act as if they do not care anymore about their information privacy or as if they value it on a different basis than the sole protection of their intimacy over the internet. What constitutes an ideal privacy as a social norm under the context of web 2.0? What is the actual impact of an improved preservation of privacy on individuals' welfare if they do not care about it? How can a balance be found between (more certain) welfare gains in terms of innovation and (more uncertain) welfare losses in terms of a reduction of privacy? To some extent, government should soften the legal standards for the protection of personal data in order to attract more internet firms and stimulate innovations in online services. A major political challenge is then to be able to rank national priorities including privacy and competitiveness.

Finally, there are limitations to this research.

First, our study is based on cross-sectional data and therefore it proves too static regarding the actual mobility of firms. One extension of this paper will consist in making a more dynamic analysis by collecting data on previous locations of firms as well as revisions to national privacy laws. Such a study would be all the more relevant since "the 'race to the bottom' could be a theoretical curiosum" if no evidence is produced about both firm mobility in response to regulatory differences and competition between states towards more lax regulation (Bhagawati & Hudec, 1995).

Second, our sample is restricted to the first hundred internet firms in the Alexa.com ranking and excludes businesses such as porn websites and B2B platforms. While it covers the major internet firms and therefore those more likely to locate overseas, it might not be representative of all internet firms. Another extension will consist in increasing our sample.

Third, as already mentioned in section 4.2, the sample consists of location decisions in terms of headquarters and does not include more detailed features of these locations, i.e. servers, marketing departments, and so on. One extension of this paper will consist in featuring the types of each location more precisely.

Lastly, we use sector-level, aggregated data – i.e. the distribution of firms according to their activities. Thus this analysis provides only an indirect test of the *individual* decisions in terms of locations. The next step will be to focus on firm-level data.

6. References

Acquisti A. (2010), "The Economics of Personal Data and the Economics of Privacy", OECD Roundtable on the Economics of Privacy and Personal Data, Paris, December 2010.

Acquisti A. (2004), "Privacy in Electronic Commerce and the Economics of Immediate Gratification", *Proceedings of ACM Electronic Commerce Conference*, New York: ACM Press, pp.21-29.

Bhagwati, J., T.N. Srinivasan (1995), "Trade and the Environment: Does Environmental Diversity Detract from the Case for Free Trade?", In J. Bhagwati, R.E. Hudec (eds.) *Fair Trade and Harmonization: Prerequisites for Free Trade?, Vol. I: Economic Analysis*, 159-223, MIT Press.

Brunnermeier S.B., A. Levinson (2004), "Examining the Evidence on Environmental Regulations and Industry Location", *Journal of Environment and Development*, vol.13, n°1, pp. 6-41.

Cairncross F. (1997), *The Death of Distance*, Boston, MA: Harvard Business School Press.

Choi, Y.J., C.S. Suh (2005), "The death of physical distance: An economic analysis of the emergence of electronic marketplaces", Papers in Regional Science, 84(4): 597-614.

Drezner D.W. (2006), "The race to the bottom hypothesis: An empirical an theoretical review", mimeo, The Fletcher School Tufts University.

Feinberg, S.E., M.P. Keane (2003), "Accounting for the Growth of MNC-Based Trade Using a Structural Model of U.S. MNCs", Yale University, Mimeo.

Fujita M., J.-F. Thisse (2002), *Economics of Agglomeration*, Cambridge University Press.

Fujita M., P. Krugman, A.J. Venables (2001), *The Spatial Economy: Cities, Regions, and International Trade*, MIT Press, Cambridge, Massachusetts.

Goldfarb A., C. Tucker (2011), "Privacy regulation and online advertising", *Management Science,* 57(1): 57–71

Grossman, G. M., E. Helpman, A. Szeidl (2006), "Optimal Integration Strategies for the Multinational Firm", *Journal of International Economics*, 70: 216-238.

Guimaraes P., O. Figueiredo, D. Woodward (2003), "A Tractable Approach to the Firm Location Decision Problem", *Review of Economic and Statistics*, 85(1): 201-204.

Hahn R.W., A. Layne-Farrar (2001), *The benefits and costs of online privacy legislation*, document de travail 01-14, AEI-Brooking Joint Center for Regulatory Studies.

Hahn R.W. (2001), "An assessment of the costs of proposed online privacy legislation", rapport, AEI-Brookings Joint Center for Regulatory Studies, Washington, DC.

Henisz WJ. (2000), "The institutional environment for multinational investment". *Journal of Law Economics and Organization*, vol.16, pp.334-364.

Hermalin B.E., M. Katz (2006), "Privacy, property rights and efficiency: the economics of privacy as secrecy", *Quantitative Marketing and Economics*, vol.4, n°3, pp.209-239

Jones G, H. Davis (2000), "National culture and innovation: implications for locating global R&D operations", *Management International Review*, vol.40, pp.11-39.

Levchenko A.A. (2007). "Institutional Quality and International Trade". *Review of Economic Studies,* 74: 791–819.

McFadden D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior" (pp.105-142), in P. Zarembka (Ed.), *Frontiers in Econometrics*, New York: Academic Press.

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

Porter M.E. (1998), Clusters and the new economics of competition, *Harvard Business Review,* 76, 6, 77-90.

Posner R.A. (1981), "The economics of privacy", *American Economic Review*, vol.71, n°2, pp.405-409.

Privacy International (2007), *National Privacy Ranking 2007 - Leading Surveillance Societies Around the World*, available at https://www.privacyinternational.org.

Stigler G.J. (1980), "An introduction to privacy in economics and politics", *Journal of Legal Studies*, vol.9, n°4, pp.623-644.

Swire P.P (1997), "Markets, self-regulation, and government enforcement in the protection of personal information", *in Privacy and Self-Regulation in the Information Age*, U.S. Department of Commerce, Washington, D.C.

Tang Z., Y. Hu, M.D. Smith (2008), "Gaining Trust Through Online Privacy Protection: Self Regulation, Mandatory Standards, or Caveat Emptor", *Journal of Management Information Systems,* vol.24, n°4, pp.152-173.

Tole L, G. Koop (2008), "Do Environmental Regulations Affect the Location Decisions of Multinational Gold Mining Firms? ", Mimeo, University of Strathclyde.

Venables A.J. (2001), "Geography and international inequalities: the impact of new technologies", *Journal of Industry, Competition, and Trade*, n°1, pp.135-160.

Wooldridge, J.M. (2001), "Econometric Analysis of Cross Section and Panel Data," MIT Press Books, The MIT Press, edition 1, volume 1, number 0262232197, June.

Yeaple S.R. (2003), "The Complex Integration Strategies of Multinationals and Cross Country Dependencies in the Structure of Foreign Direct Investment". *Journal of International Economics*, 60(2): 293-314.

Appendix

Table A1: Definition of Variables

Variable	Definition	Source
Privacy Index	The index is a scoring calculated with thirteen criteria: Constitutional protections; Statutory protection; Privacy enforcement; Identity cards & biometrics; Data sharing provisions; Visual surveillance; Communications interception; Workplace monitoring; Law enforcement access; Data retention practices; Travel & finance surveillance (including trans-border data sharing); Global leadership; Democratic safeguards	Privacy International (*The Privacy & Human Rights Report*, 2007)
Internet Users	"people with access to the worldwide network."	World Bank Development indicators + ITU
Profit Tax	"the amount of taxes on profits paid by the business." (% of commercial profits)	World Bank Development indicators
Political Stability	Composite index (from 1 to 4) built from: - "Have the rules in force when the current head of state or government came to power changed in the last three years to improve the head of state or government's chances of staying in his/her position?" (from 1: rules significantly changed in 3 years to 4 = unchanged) - "Did the current head of state or government come to power in compliance with the (changed) rules in force at the moment of his/her accession (election, dynastic succession, etc.)?" (from 1 = massive breach of the rules in force to 4 = total compliance with the rules) The legality of the process appointing these political authorities (from 1 = weak legal to 4 = high legal process)	Cepii
Transparency of Public Economic Actions	Transparency of economic policy (fiscal, budgetary, monetary, exchange rate, etc.) is a composite index (from 1 to 4) built from: - "Is economic policy subject to an official communication? (from 1= thin and opaque to 4= substantial and transparent)" -"Is economic policy subject to public debate? (from 1= little debate to 4 = substantial and structured debate)"	Cepii
Business Technological Environment	Technological environment and dissemination of technology are evaluated through a composite index (from 1 to 4) built from: "Proportion of technical staff (engineers, technicians) in SMEs/SMIs (from 1 = low to 4 = high)"; "Proportion of technical staff (engineers, technicians) in large firms (from 1 = low to 4 = high)" ; "Cooperation links between businesses, universities and research centers (from 1 = few links to 4 = close links)"	Cepii
Number of Servers	Number of Internet hosts available within a country. An Internet host is a computer connected directly to the Internet: normally an Internet Service Provider's (ISP) computer is a host. Internet users may use either a hard-wired terminal, at an institution with a mainframe computer connected directly to the Internet, or may connect remotely by way of a modem via telephone line, cable, or satellite to the Internet Service Provider's host computer. The number of hosts is one indicator of the extent of Internet connectivity.	CIA World Factbook
Quantity of Civil Societies	Non-governmental organization that "gives access to intergovernmental process at the UNO dealing with economic and social development, gender issues, sustainable development, small arms, and human rights".	United Nations Department of Economic and Social Affairs
Freedom of the Press	This classification is established by Reporters Sans Frontières network which includes 15 associations for the freedom of speech and expression, 130 correspondents, journalists, researchers, lawyers, or human rights militants. This network answers 50 questions allowing the evaluation of the freedom of the press situation in each country.	Reporters Sans Frontières

Table A2. The construction of the dependent variable:

The dependent variable is the sum of internet firms (in our sample) in a given sector established in each country. For example, Argentina hosts one firm in sector one, one firm in sector five and two firms in sector 6. The privacy index for Argentina is 2.8.

Countries / Sectors	1	2	3	4	5	6	7	8	9	10	***Privacy index***
ARE	1	1				1			1		-
ARG	1				1	2					*2,8*
AUS	3	1			2	1			1		*2,3*
…											
USA	38	20	24	8	15	27	5	1	6	1	*1,5*
VNM						1					-
ZAF						2			1		2,3

Description of the sectors		Web 2 sectors
1	search engines & web portals	-
2	online networking services	√
3	web hosting & sharing services	√
4	online press & information services	-
5	infomediaries	-
6	e-commerce	-
7	corporate websites	-
8	online administration	-
9	e-advertising	-
10	online games	-

A3. Statistical distribution of the dependent variable

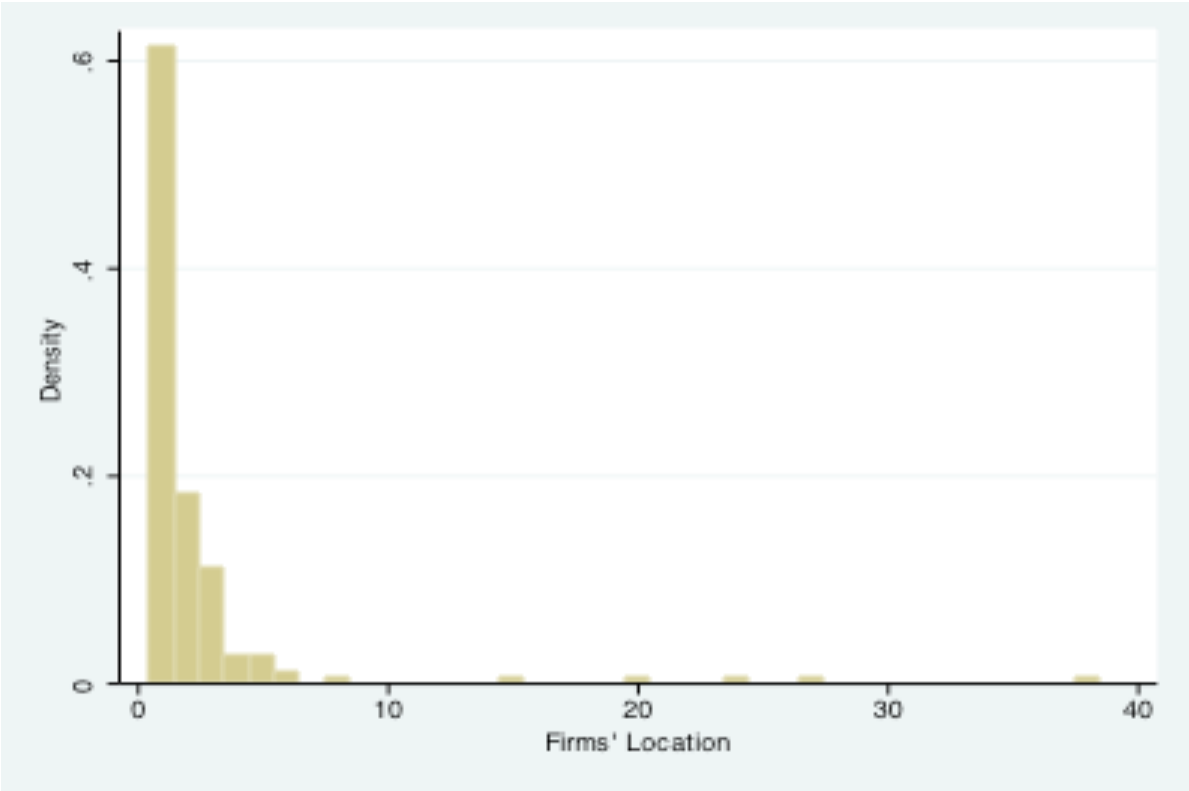

Table A4 – Econometric Results – Negative Binominal Regressor

Dependent variable: Firms' Location	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	All	All	All	All	No USA
Privacy	-0.89*	-0.89*	-1.15**	-1.23***	-0.67***	-0.36**
	(0.46)	(0.46)	(0.45)	(0.33)	(0.14)	(0.16)
Ln(Tax)		-0.03	0.14	0.70*	-0.45	-0.09
		(0.28)	(0.30)	(0.38)	(0.30)	(0.26)
Political Stability			1.74***	1.25***	0.67**	0.49**
			(0.65)	(0.38)	(0.29)	(0.23)
Transparency of public economic actions				0.03	0.25	0.1
				(0.16)	(0.15)	(0.10)
Business Technological Environment				0.67***	0.22**	0.06
				(0.23)	(0.11)	(0.10)
Internet Users					-0.47*	0
					(0.28)	(0.21)
Ln(Servers)					0.38***	0.13**
					(0.04)	(0.05)
Constant	3.09***	3.19**	-3.75	-6.14***	-5.64***	-2.52**
	(1.17)	(1.58)	(2.30)	(2.29)	(1.33)	(1.15)
Observations	152	152	152	152	152	142

As probchibar2<alpha for the last regression, a negative binomial estimator can provide robustness. Nonetheless, results for privacy are very close to those of the Poisson QMLE estimator. Robust standard errors in brackets are clustered by country. *** p<0.01, ** p<0.05, * p<0.1. All regressions include sector fixed-effects.

Table A5: Spillovers Robustness Test

Dependent variable: Firms' Location	(1)	(2)	(3)
Sample	All	All	No USA
Privacy	-0.60***	-0.41***	-0.19**
	(0.17)	(0.13)	(0.10)
Ln(Tax)	-0.42	-0.19	-0.16
	(0.31)	(0.22)	(0.17)
Political Stability	0.63	0.52*	0.3
	(0.39)	(0.26)	(0.20)
Transparency of Public Economic Actions	0.12	0.12	-0.04
	(0.17)	(0.12)	(0.09)
Business Technological Environment	0.19	0.1	0.09
	(0.14)	(0.10)	(0.08)
Internet Users	-0.26	-0.12	-0.18
	(0.42)	(0.30)	(0.23)
Ln(Servers)	0.33***	0.18***	0.02
	(0.05)	(0.06)	(0.05)
Other Internet Firms		4.68***	22.70***
		(1.17)	(3.47)
Web 2 Dummy	0.19		
	(0.15)		
Constant	-5.11***	-2.80**	0.01
	(1.61)	(1.29)	(1.04)
Observations	152	152	142
Sector fixed effects	No	Yes	Yes

Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. As data presents overdispersion, standard errors are corrected by the Pearson chi-squared statistic divided by the degrees of freedom.

Table A6 – Poisson QMLE Intrumental Variables Including Business Model 4 (online press)

Dependent variable: Firms' Location	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	No USA	All	No USA	All	No USA
IV	Civil soc.	Civil soc.	Press freedom	Press freedom	Both	Both
Privacy	-1.16***	-0.76***	-1.04***	-0.65***	-1.08***	-0.75***
	(0.17)	(0.18)	(0.18)	(0.16)	(0.13)	(0.15)
Tax	-0.05	0.02	-0.39*	-0.06	-0.16	-0.02
	(0.25)	(0.21)	(0.22)	(0.20)	(0.23)	(0.21)
Political Stability	1.11***	0.79***	0.97***	0.73***	1.07***	0.80***
	(0.20)	(0.18)	(0.20)	(0.18)	(0.18)	(0.17)
Transparency of Public Economic Actions	0.20*	0.13	0.29***	0.13	0.21**	0.14
	(0.11)	(0.09)	(0.10)	(0.09)	(0.09)	(0.09)
Business Technological Environment	0.26***	0.16*	0.31***	0.14	0.28***	0.16**
	(0.10)	(0.08)	(0.10)	(0.08)	(0.09)	(0.08)
Internet Users	-0.36	-0.13	-0.60**	-0.13	-0.41*	-0.16
	(0.25)	(0.20)	(0.25)	(0.21)	(0.23)	(0.20)
Ln(Servers)	0.27***	0.16***	0.38***	0.16***	0.29***	0.16***
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)
Constant	-6.01***	-3.92***	-6.59***	-3.57***	-6.15***	-3.97***
	(0.95)	(1.07)	(1.01)	(1.05)	(0.91)	(1.03)
Observations	152	142	152	142	152	142
Wu-Hausman (p-value)	0.0000	0.0001	0.0028	0.0022	0.0000	0.0018
Sargan/Hansen (p-value)					0.8930	0.7812

Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. As data presents overdispersion, standard errors are corrected by the Pearson chi-squared statistic divided by the degrees of freedom. All regressions include sector fixed-effects. See first step in table A8 below.

Table A7 – First step of Table 3 (Instrumental variables)

Dependent variable: Firms' Location	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	No USA	All	No USA	All	No USA
IV	Civil soc.	Civil soc.	Press freedom	Press freedom	Both	Both
Civil societies	-0.62***	-0.97***			-0.42***	-0.51***
x 1000	(0.08)	(0.16)			(0.08)	(0.17)
Freedom of the Press			-0.03***	-0.03***	-0.02***	-0.02***
			(0.00)	(0.00)	(0.00)	(0.00)
Tax	0.17	0.19	0.44***	0.25**	0.23**	0.25**
	(0.14)	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)
Political Stability	0.79***	0.87***	0.09	0.18	0.30**	0.31**
	(0.14)	(0.13)	(0.15)	(0.15)	(0.14)	(0.15)
Transparency of Public Economic Actions	0.21***	0.21***	-0.36***	-0.31***	-0.18**	-0.16
	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)	(0.10)
Business Technological Environment	0.24***	0.42***	0.22***	0.25***	0.24***	0.24***
	(0.06)	(0.07)	(0.06)	(0.06)	(0.05)	(0.06)
Internet Users	-0.79***	-0.69***	-1.09***	-1.15***	-1.27***	-1.30***
	(0.21)	(0.21)	(0.21)	(0.20)	(0.19)	(0.20)
Ln(Servers)	0.12***	0.14***	-0.14***	-0.04	0.01	0.01
	(0.04)	(0.04)	(0.02)	(0.04)	(0.04)	(0.04)
Constant	-4.22***	-5.12***	4.08***	2.60***	1.08	0.91
	(0.77)	(0.75)	(0.92)	(0.96)	(1.00)	(1.08)
Observations	143	134	143	134	143	134
R-squared	0.52	0.55	0.57	0.58	0.65	0.61

Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. All regressions include sector fixed-effects.

Table A8 – First step of Table A6 (instrumental variables)

Dependent variable: Firms' Location	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	No USA	All	No USA	All	No USA
IV	Civil soc.	Civil soc.	Press freedom	Press freedom	Both	Both
Civil Societies	-0.63***	-0.94***			-0.44***	-0.39**
x 1000	(0.08)	(0.17)			(0.07)	(0.16)
Freedom of the Press			-0.03***	-0.03***	-0.02***	-0.01***
			(0.00)	(0.00)	(0.00)	(0.00)
Tax	0.16	0.22	0.45***	0.24*	0.23**	0.32***
	(0.13)	(0.14)	(0.12)	(0.12)	(0.11)	(0.12)
Political Stability	0.80***	0.81***	0.08	0.16	0.29**	0.49***
	(0.14)	(0.14)	(0.15)	(0.15)	(0.14)	(0.15)
Transparency of Public Economic Actions	0.23***	0.28***	-0.38***	-0.33***	-0.18**	-0.20**
	(0.07)	(0.08)	(0.09)	(0.08)	(0.09)	(0.10)
Business Technological Environment	0.24***	0.23***	0.21***	0.24***	0.24***	0.34***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.07)
Internet Users	-0.77***	-0.90***	-1.03***	-1.11***	-1.25***	
	(0.20)	(0.22)	(0.20)	(0.19)	(0.18)	
Ln(Servers)	0.13***	0.12***	-0.15***	-0.03	0.02	0.02
	(0.04)	(0.04)	(0.02)	(0.03)	(0.04)	(0.04)
Constant	-4.38***	-4.57***	4.24***	2.64***	1.04	-0.85
	(0.76)	(0.78)	(0.92)	(0.95)	(0.99)	(1.14)
Observations	152	142	152	142	152	142
R-squared	0.53	0.48	0.57	0.58	0.66	0.58

Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. All regressions include sector fixed-effects.