EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Charles Hoffreumon, Vincent Labhard

Cross-country cross-technology digitalisation: a Bayesian hierarchical model perspective

No 2700 / August 2022

Abstract

In this article, we present a new perspective on forecasting technology adoption, focused on the extensive margin of adoption of multiple digital technologies in multiple countries. We do this by applying a Bayesian hierarchical structure to the seminal model of technology diffusion. After motivating the new perspective and the choices of priors, we apply the resulting framework to a cross-continental data set for EU and OECD countries and different digital technologies adopted by either households/individuals or by businesses. The results illustrate that the Bayesian hierarchical structure may be used to assess and predict both the adoption process and the uncertainty surrounding the data, and is robust to the use of alternative priors. They point to heterogeneity across countries and across technologies, mostly in the timing of adoption and, although to a lesser extent, the steady-state adoption rate once technologies are fully diffused. This suggests that characteristics of countries and technologies matter for technology diffusion.

Keywords: Adoption, Diffusion, Timing, Speed, Maximum

JEL codes: C11, C52, C53, O33, O57

Non-Technical Summary

Technology in general and digital technology in particular, is an important component of everyday lives, and therefore important also from a policy perspective. Robots and automation have changed how products are being made, and the internet of things electronically connects appliances. Computers and mobiles have changed what people consume and how they communicate, and the internet serves as a source of information, a medium of exchange, and as an enabler for many types of transactions.

In this context, a key concern is the ability to understand what is or what will be the bundle of technologies available. In relation to households/individuals, the question is whether and when they are going to have access to enabling technologies, as for the example email and the internet, and the technologies building on those, such as social networks. As regards businesses, the same question arises regarding broadband, enterprise resource planning, e-commerce and customer relationship management, among others.

The specific technology being considered is one factor determining the process of technology diffusion. The literature has also considered a number of other factors, such as, most recently, the influence of institutions and governance. If institutions and governance are of high quality, it may be easier to adopt digital technology, for example to get access to the internet, or use search engines, or to use a mobile phone, or a robot. Given that institutions and governance are specific to countries, the process of technology diffusion is also going to be specific to specific countries.

This paper approaches modelling and estimating the process of technology diffusion from a new perspective, taking into account both technology-specific and country-specific factors. That new perspective is Bayesian, and its essence is that the parameters of the model are estimated starting from priors which are then contrasted to the data. In this paper, the priors relate to the timing of adoption, the speed of adoption, and the ultimate level of adoption, which all are a function of both technology and country.

The idea is applied to data for a set of 30 countries which are members of European Union or OECD, and 14 digital technologies, 9 that are particularly pertinent for households/individuals and 5 that are better thought of as related to businesses, for the period 2005-2020. The results suggest that several technologies have yet to diffuse fully, especially those related to businesses, perhaps because those tend to be more complex. There is evidence also that the timing of

adoption and, to some extent, the ultimate rate of adoption tend to differ across countries. There is evidence also that later adoption does not mean faster adoption, at least not in connection with technologies associated more to businesses.

Those results on the forecasting aspect are robust, and the diagnostics suggest that the Bayesian perspective on technology diffusion is well specified. There seems to be scope though for further enhancing the results, especially by means of additional data. The key challenge in this respect is that the technologies that are the most interesting – those that are the most recent to have been introduced and about to diffuse widely – are those for which the data tend to be most limited. One of the greatest advantages of the framework presented in this paper though is that it can provide estimates even if those data are available only to some extent.

1 Introduction

Technology in general, and digital technology in particular, is an important component of our economic lives. Generally, many aspects of our day-to-day activities have become more efficient, or even possible, by the access to a stock of technology stemming from a continuous stream of past innovations. Not all of us have the same access though, as a function for example of where we live. As such, the ability to understand what is or predict what will be the technological bundle installed in the future is important and this paper presents a new perspective for monitoring and forecasting technology adoption.

This new perspective enables working with the noisy data on technology adoption typically available and performing well even when the available data is sparse or completely missing. This may happen, for example, in the case of technologies that have been introduced most recently and so have yet to diffuse widely, and those technologies in some ways are the most interesting, because the impact they may have on the economy has yet to unfold. The ability of the framework to cope with missing data therefore is important for the task at hand.

Moreover, data on technology adoption at the country level is typically often available from surveys. Frequently, the results do not originate from the same firms or households. Additionally, data construction and coverage may vary greatly from country to country. The ability of forecasting the mean technology adoption while remaining tolerant of missing values and outliers without requiring a panel data set at the level of the firm are therefore desirable features of a framework for monitoring and forecasting technology.

Modern methodological advances, notably in the field of computational statistics, allow modelling "knowledge transfer" from one technology to the other and from one observational unit to the other, and are computationally tractable. One such computational methodology, employed in this paper, is hierarchical Bayesian modelling. Even though Bayesian inference using approximation of the full posterior (as opposed to the use of conjugate priors) has been a topic of research since the mid-twentieth century, it has recently gained both performance and popularity, mostly thanks to the advances in the machine learning field.

The framework presented here is based on a hierarchical Bayesian structure, that allows for both commonality and variability across units in a Bayesian context, but otherwise is deliberately held as simple as possible. It reuses concepts both from classical literature on technology adoption and recent progresses on statistics to be both easily interpretable and suited to the application of monitoring and forecasting the rate of technology adoption directly, and without requiring the researcher to impute missing values.

There are not many studies that aim at spanning several technologies or countries. Most research to date has been focused on predicting the adoption of specific technologies in specific industries or geographic zones. An important exception, however, are Comin and Hobijn [2010], who formalise the problem of estimating diffusion across several geographies and technologies, and then explore reasons of adoption of certain technologies in certain countries. But the scope of the data set used, both in terms of time and technologies (with each a certain unit of measure of adoption, since their paper discusses the intensive margin), makes it less well-suited to what we propose here.

The literature on technological adoption dates back several decades but has remained remarkably homogeneous in its core principles. It has its origins (inter alia) in research on agricultural innovations diffusion by Griliches [1957], later applied to technology more generally by Mansfield [1961]. The key idea in that research is that diffusion takes an "S-shape"; at the start of the process, a few "innovators" experiment with the change, and the adoption then speeds up until an inflexion point, followed by a reduction in the growth rate of adoption until, finally, adoption reaches a *plateau*.

The reasons for the emergence of such a shape at the macro level have been extensively researched in several fields of human sciences, most notably in marketing, sociology and economics, which have each identified different mechanisms at the micro level and the shape they may entail at the macro level when aggregated (an account of which is given in Young [2009]). While in some cases it seems it is possible to favour one explanation over another based on the picture at the macro level, discriminating between them would most often require access to micro data, which are unavailable for a large number of technologies.

Another paper, from the field of marketing, proposes building hierarchical models as a way to estimate the maximum adoption rate of different (related) technologies. Lenk and Rao [1990] essentially advocate using a model similar to the one examined in this paper, but apply it to a single market, using micro-level data. That paper is, however, a *tour de force* as it predates the recent advances in efficient Monte Carlo Markov Chain (henceforth MCMC) and yet manages to obtain results by exploiting conjugacy. Using today's machinery, many of the assumptions and transformations that were necessary to make the ideas proposed in the paper tractable can be lifted and we have liberty to choose amongst a broader range of priors, among other options, to make the model applicable to multiple markets.

Also closely related to this paper is the work of Lee et al. [2003]. In it, the authors build a model to predict the sales of new music albums based on sales of previous albums. The hierarchical prior is used to set the market potential for each album. In this case they do not make any attempt to segregate by markets.

By combining the insights of several strands of literature, the perspective presented here may be used to monitor technology diffusion and to make an educated guess on the likely adoption path of a specific technology in a specific geography going forward. It does so by allowing partial sharing of coefficients across zones and country and provides a set of priors that are flexible enough to adapt to a broad range of underlying processes.

Moreover, we allow for heterogeneity in yet one more way, by allowing estimation and forecasting also separately depending on who technologies are used by - households/individuals (also referred to in this paper as consumers) or businesses (alternatively labelled producers, companies or corporates) - given the theoretical underpinnings suggest they may be characterised by specific behaviours and so experience different diffusion dynamics.

The resulting framework is designed to be easily extensible to include information about the regional units or the technology, serving as a test-bed to assess potential factors that might contribute to earlier or faster adoption, as for example institutions and governance as in Baccianti et al. [2022]. We discuss some of these extensions in the latter part of the paper.

2 Framework

2.1 The seminal diffusion model from a Bayesian hierarchical perspective

For this paper, we use the mathematical foundation of the Mansfield model (Mansfield [1961]). That model, seminal in the literature on technology diffusion, is a particular case of the more widely-known extended Bass Model (Mahajan et al. [1990]) in which diffusion happens exclusively through imitation. Despite its early publication date, this model is very much in use in forecasting adoption of new products and technology (for instance, Stoneman [2011], Sudtasan and Mitomo [2017], Jha and Saha [2018]).

Conceptually, in the Bayesian hierarchical perspective taken here, the diffusion process has two types of entities: technologies, for which a distinction is going to be made between technologies adopted by consumers (households/individuals) and those adopted by producers (businesses), and countries. Both entities have particular features that are intrinsic to them.

Let's first consider technologies. Technologies are invented, or start being popularised, at some period in time, some of them are "easier", or faster, to adopt than others, and they are not all relevant for the whole population. While the timing of diffusion merely represents the point in time at which diffusion reached or will reach its inflexion point, the rate at which they are adopted and the share of the population that is ultimately going to adopt them might merit further discussion. We take them in turn.

The "intrinsic speed" of diffusion of a technology might arise from several factors that have been researched in the literature. This is, in fact, the object of Young [2009], who identifies three main potential sources of diffusion. The first is "contagion", the second is social imitation and the third is social learning. Out of those, he singles out social learning as the most likely explanation. In other word, individuals and companies adopt technologies once they have enough evidence that said technology will offer benefits to them. Following this, there is no reason to think that different technologies, each with their own characteristics, will have the same pattern of diffusion. This is why judgemental forecasts of technology diffusion that one can occasionally find in the media are usually to be taken with a grain of salt. They depend, often substantially, on the belief of the person formulating them of how close an innovation is to previous innovations. In our framework, we do away with such an a-priori judgement by assuming that completely unknown technologies are some sort of combinations of all past technologies in the set and then, as we gain more information on it, we update our parameter estimates accordingly.

The maximum adoption likewise originates from the rationale of social learning. Since consumers and producers adopt technologies that they find beneficial for them, there is no reason to believe that technologies are, if adopted, beneficial for the whole population or for a fixed share thereof. As such, we estimate this criterion but do not attempt, when confronted with a completely unheard-of technology, to consider it as a mix of all other. Rather, we start with the initial assumption that technologies that are measured, because they made their way into the set of technologies tracked by statistical bureaus, will usually reach a significant part of the population, but we remain extremely vague about the exact quantity and let the "data speak" relatively early on.

Next, let's consider the other conceptual entity in the model, the countries. We see countries as being heterogeneously "permeable" to technologies on two aspects. On the one hand, we let the model allow that a country adopts a certain technology earlier than its peers. The second aspect we model is that, initially independently from this "lag" in the timing of adoption, some countries, once they are on the path to adoption, reach the full adoption of the technology faster, ending the diffusion process sooner than another country that would have the same "lag" but a smaller adoption speed modifier coefficient.

While simple, with only technologies and countries to consider, and the diffusion process fully characterised by three parameters, the framework used here is surprisingly well-suited to represent the diffusion process, provided one does not enforce too restrictive priors. Indeed, as will be seen in the next section, and while not making any claim at describing the causal process of adoption, this model can be used to represent technology diffusion and the uncertainty arising from both the process itself and the measurement.

2.2 Why a Bayesian Hierarchical Model Perspective?

The framework detailed here enhances the scope of the Mansfield model in a number of ways. While Mansfield [1961] and other authors have traditionally sought to estimate the mean adoption and then used the model to assess the impact of different events or characteristics on this mean adoption, the present model seeks to estimate a probability density of adoption lying in a certain span, given the data already at hand.

The different objective has two main implications:

1. The estimation needs to start from prior distribution and make those priors wide enough so that they do not have too much impact on the final distributions.

2. The assessment of the model should be done in such a way that we do not only check the alignment of the means but of all the quantiles of the distributions, forecast and observed.

Rather than an entirely new model with new functional form of the mean, it is an extension where, besides the mean, we seek to estimate the variance and correlations between countries and between technologies. This is done with the idea of being able to incorporate the notion of risk into potential higher-level models. By being more flexible with the error terms and incorporating the results of recent research in both statistics (Ferrari and Cribari-Neto [2004]) and computer science (Hoffman and Gelman [2014]), we provide a tool that makes forecasting adoption able to fit future data more closely while remaining feasible with relatively few data from the past.

2.3 The Probability Density of the Rate of Adoption

The specification of the model as outlined above, with the focus on the diffusion among populations of firms or households/individuals, has two important consequences:

1. The diffusion process is binary
2. The image (the domain of definition of the result) of the diffusion function is constrained

The first point is similar to most of the diffusion models, including the canonical Bass model. It means that members of the population either adopt the technology or do not. In the technology adoption literature, this is called the extensive margin of adoption. Estimating the so-called intensive margin of adoption (for instance, not only the share of the population that uses internet banking but also the share of banking transactions performed via the internet) is not the original purpose of the model but can be accommodated if the intensive margin is expressed as percentage of a theoretical maximum. This point, however, might benefit from further research.

The second point orients the choice of the likelihood or probability density function. As detailed in Ferrari and Cribari-Neto [2004], a sensible choice in such situations is the Beta distribution. This distribution is appropriate where the variable of interest is a proportion or a rate (which is the case with the adoption rate). The probability density for the adoption rate is expressed using the alternative parameterisation of the Beta distribution (also found in Ferrari and Cribari-Neto [2004]):

$$\pi(y_{ijt}; \mu_{ijt}, \epsilon_i) = \frac{\Gamma(\epsilon_i)}{\Gamma(\mu_{ijt}\epsilon_i)\Gamma((1-\mu_{ijt})\epsilon_i)} y_{ijt}^{\mu_{ijt}\epsilon_i - 1}(1-y_{ijt})^{(1-\mu_{ijt})\epsilon_i - 1} \tag{1}$$

where the outcome variable, y_{ijt} is the share of the sample (be it consumers or companies/firms) in country i who declared to have adopted technology j in year t, ϵ_i is the concentration term[1] associated with country i and Γ is the gamma function. The next subsection, 2.4, details the specification of the μ and ϵ coefficients.

[1]The concentration term is a term associated with the inverse of the variance of the distribution. Assuming the variance of y to be $var(y)$, the concentration term is $\epsilon = \frac{\mu(\mu-1)-1}{var(y)}$. A higher concentration term is therefore associated with a lower variance and priors must therefore be chosen to be sufficiently low rather than high when estimating the model to reflect uncertainty on the parameters.

2.4 The Mean of the Likelihood

As for most regression models, most of the attention needs to be paid to the mean of the outcome. For the purpose of this paper, we want to capture two sets of information: information about the country or geographical entity in which adoption takes place (hence, enabling the incorporation of economic, geographical and institutional conditions). In this paper, we won't try to represent any of the characteristics of the country or technology, but use fixed effects for these dimensions. The addition of such characteristics might be the focus of further extensions of the framework considered here.

The sigmoid function we choose to represent the adoption process is the classical logistic function from Richards [1959]. The function is defined, for country i and technology j, as:

$$\mu_{ij}(t) = \frac{o_{ij}}{\left(1 + e^{-\nu_{ij}(t - t_{0ij})}\right)^{\frac{1}{\omega_i}}} \tag{2}$$

with four parameters:

1. o, the maximum adoption rate once the technology is fully diffused; this is estimated at the level of country and technology.
2. ν, the speed of adoption of a technology in a country; we estimate one per technology/country pair.
3. t_0, the timing of the adoption of a technology in a country; it is similarly defined at the technology/country pair level.
4. ω, the shape parameter that determines the behavior of the curve near the asymptotes (a value for this parameter of 1 corresponds to a symmetrical curve); this parameter is estimated at the level of the technology.

2.4.1 Maximum adoption (o)

The maximum rate of adoption is allowed to vary according to the country and technology. It is computed as the logistic transform of the sum of country and technology intercepts, denoted below by χ_i and κ_j respectively.

$$\hat{o}_{ij} = \frac{1}{1 + e^{-(\chi_i + \kappa_j)}}$$

The priors on both coefficients are standard normal distributions.

2.4.2 Speed of Adoption (ν)

The speed of diffusion depends on both country and technology. We assign the same type of prior to both the components initially although, as we will see in Section 3, the effect of technology is much more important than the peculiarities of the countries. The speed of diffusion of technology j in country i is therefore computed as:

$$\hat{\nu}_{ij} = \delta_i + \beta_j$$

where β_j represents the "intrinsic speed" at which the focal technology diffuses and δ_i a modifier of the speed of diffusion of a technology dependent on the country.

The priors, listed in Table A.5 in Appendix A, aim at representing the fact that the diffusion of a technology will, in most cases, take a few periods (years) before reaching its maximum value. We do not, however, impose strict limits as a consequence of these priors. Indeed, both for the country effect (δ_i) and the technology effect (β_j), we use a hierarchical prior on both the mean and the variance terms of these coefficients.

Inserting such a hierarchical structure makes the inclusion of new technologies in the model easier. Indeed, new technologies will likely enter the model with very little information. Being able to tap into the value of the coefficients even for technologies for which we have very little information is an asset of the model presented here. The use of a hierarchical prior at the level of the technology has, of course, the effect of bringing the mean of the adoption speeds closer together but this drawback allows for better estimates even for new, previously unseen technologies to be more accurate considering what is known about other technologies. This idea is similar in nature to the one in Lenk and Rao [1990] and the interested reader can find there an in-depth discussion of the mechanism.

2.4.3 Timing of Diffusion (t_0)

The estimation of the t_{0ij} is also the sum of two components, defined at the technology and country levels. The major difference with the speed of adoption is that, in this case, we do not allow partial pooling of the mean coefficients at the level of the technology. Indeed, each technology is unique in its invention date and the time it takes to start being adopted. As such, knowing when a certain invention was invented or became popular does not “teach” anything about when another technology was invented or started being popular. This independence is partly due to the simple nature of the fixed effects of the model and may be lifted as one adds lower-level explanatory variables. Formally,

$$\hat{t}_{0ij} = \theta_i + \tau_j$$

where θ_i is the modifier specific to country i and τ_j is the timing intrinsic to the technology j. In contrast to the other parameters, the parameters θ_i and τ_j have natural units, the years of delay relative to the start of the data set (in 2005).

The wide hierarchical prior chosen to accommodate the recency of the technologies we set out to analyse in Section 3 turns it rather uninformative. However, it remains a bit awkward not to provide any information about the timing of adoption. While this may seem like a weakness, it stems from a choice to keep the structure very general at this stage. Indeed, specifying the timing of adoption relative to the time a technology reaches a certain threshold, as in Comin and Hobijn [2010], requires a measure of the adoption at that point which makes the framework hard to use for very recent technology. By keeping the timing coefficient floating, the present framework allows providing an own prior for this parameter depending on the purpose of the inference.

2.5 Concentration Term (ϵ)

Finally, the error is estimated at the level of the country. This choice is justified by that, under the beta regression specification, the error (or, more specifically, the concentration term, which is related to the error as explained in footnote 1) represents the “number of respondents” in an hypothetical survey that would have generated the result. These coefficients represent the “degree of confidence” one might have that the value is close to the mean.

The choice of pooling the error at the level of the country therefore follows the rationale that countries with different size will likely have different sample size, as the data originates from a compilation of different country-level surveys operationalised by the different national statistical offices.

2.6 The choice of priors

The priors were chosen so as to be wide enough to let the data "speak for itself" while allowing pooling to perform the out-of-sample estimates. They are based (with only two exceptions) on Normal or HalfNormal (and, with only one exception, normalised) distributions, namely Uniform(0.0,1.5) for the shape parameter ω_j, Normal(0.0,1.0) for maximum adoption, and for speed ($\nu_i j$) and timing (t_0) of adoption (for the hierarchical mean parameters, HalfNormal(5.0) for the hierarchical variance parameters).[2] To avoid assuming too precise a model, we chose a Gamma(5.0,1.0) prior for the error term (concentration) ϵ_i, as this prior has a sizeable mass near the origin.

The values of the priors are documented also in Table A.5 in Appendix A. For the interested reader, a detailed representation of the model including the information on the priors in the form of a full directed factor graph is given in Figure B.1 in Appendix B.

3 Data and Estimation

We start by discussing the characteristics of both data sets in the next section and then briefly turn, without delving into details, to the inference method. We then analyse the results, highlighting points where work remains to be done.

3.1 Data

We use our Bayesian hierarchical structure on two data sets. Both are from the OECD and originate from their series of surveys on information and communication technologies (ICT). The two data sets, the "ICT Access and Usage by Households and Individuals" and the "ICT Access and Use by Businesses" are freely available. Both are, in fact, an aggregation of recurring surveys performed by national statistical agencies. In most OECD countries, such surveys are

[2] The robustness of the results to the choice of priors has been monitored closely during the production of the paper, but mostly in the context of changes to both model and priors in parallel. A more exhaustive and systematic analysis of this aspect of the paper is planned for the journal version of the paper.

Table 1: Descriptive statistics on the technology adoption by households/individuals

Technology	obs (#)	ctrs (#)	first (year)	last (year)	min (%)	max (%)	most (pp)	least (pp)
Internet	452	29	2005	2020	13.93	99.47	63.74	13.63
Internet Banking	465	30	2005	2020	0.39	95.93	64.51	11.41
Social Networks	277	30	2005	2020	2.90	93.82	66.50	7.26
Private Emails	441	30	2005	2020	8.68	95.97	56.22	19.69
News and Magazines	428	30	2005	2020	0.43	95.86	70.23	30.42
Health Information	394	29	2005	2020	1.12	77.15	58.63	26.35
P2P eCommerce	432	30	2005	2020	0.15	47.67	34.08	3.56
VoIP/Teleconferencing	444	30	2005	2020	0.73	82.72	77.71	37.30

Source: OECD (2020), authors' calculations.
Note: 'P2P' stands for 'peer-to-peer', 'VoIP' for 'voice-over-IP', 'Health Information' for 'Health Information on the Internet'. The column 'most' ('least') displays the difference in percentage points between the maximum and minimum adoption for the country where that difference is the most (least).

not part of the core surveys run every year (such as the large scale labour force surveys in the EU or the census surveys in the US). They are often carried out on an ad-hoc basis, sometimes in order to complement a larger survey and with a varying set of questions. This has implications for the analysis of the results over several years.

While Bayesian methods are notably resilient to irregular and missing data, extreme scarcity of data could make the estimation phase of the analysis complicated.[3] For this reason, we operated a selection to guarantee that, for each technology/country pair, we have at least 9 data points. This is of course only a modelling choice, and this criterion can be loosened at the cost of larger uncertainty on the coefficients and results.

Table 2: Descriptive statistics on the technology adoption by businesses

Technology	obs (#)	ctrs (#)	first (year)	last (year)	min (%)	max (%)	most (pp)	least (pp)
eCommerce	233	29	2008	2019	4.70	34.69	14.71	0.00
CRM	189	28	2006	2019	7.29	56.10	34.25	3.63
ERP	236	28	2006	2019	4.27	56.48	43.10	3.58
Social Media	164	29	2011	2019	12.20	78.65	39.82	0.00
Broadband	219	26	2011	2019	1.94	62.15	52.59	0.93

Source: OECD (2020), authors' calculations.
Note: 'CRM' stands for 'Customer Relationship Management', 'ERP' for 'Enterprise Resource Planning'. The column 'most' ('least') represents the difference in percentage points between the maximum and minimum adoption for the country where that difference is the most (least).

[3]This does not mean that inference will be impossible for technologies left out of the present analysis because, as was discussed in Section 2, we set and estimate hierarchical priors for the speed of diffusion at the level of the technologies.

As one can see, there is a certain degree of heterogeneity between the adoption-by-households/individuals and adoption-by-businesses data sets. The surveys on technology use by individuals and households date back further and the series span a longer time accordingly. The technologies usually end up being adopted by large swathes of the population *during the course of the data set*. On the other hand, we have shorter series in the businesses surveys and the maximum adoption of some of the technologies remains relatively low. Whether that is because technologies are still evolving towards their maximum or rather because they already reached their maximum adoption is left to be estimated.

3.2 Estimation Process

We estimate the coefficients using the No U-Turn Sampler (NUTS) from Hoffman and Gelman [2014]. This sampler is a Hamiltonian Monte Carlo algorithm and allows to efficiently sample from the posterior in a large set of situations. It requires little configuration and is used in many recent Bayesian data analyses. We let 5 parallel chains run for $10,000$ warm-up steps, followed by the collection of $10,000$ samples. This is done in order to collect enough sample to do robust inference, and is given as a rule-of-thumb in Kruschke [2010].[4] The inference steps take around 4 hours on each dataset using Numpyro (Phan et al. [2019], Bingham et al. [2018]).[5]

The diagnostic is done through the verification that none of the coefficient has a split-$\hat{R}$ above 1.05 and there are no diverging samples. This provides evidence that the chains have mixed well and that they might have reached the equilibrium distribution of states. Indeed, while there is no definitive test allowing to exclude that some important zone of the parameter space has not been explored, these measures (the split-$\hat{R}$ and, specifically for the NUTS algorithm, the absence of diverging samples) are generally used as the leading indicators for correct mixing of the chains. The interested reader is invited to consult any recent book on Bayesian data analysis (such as Gelman et al. [2013]) for reference on these concepts.

[4]The rule of thumb concerns the number of effective samples but this setup allows collecting enough samples to reach this effective sample size.

[5]An alternative way of performing the estimation of the model would be to use Approximate Bayesian Computation, such as Stochastic Variational Inference. This technique was applied with success for this model but we present the results obtained through MCMC as they have the additional property that they converge asymptotically to the true posterior, at a cost of greater computational expense.

3.3 Estimation Results

In this section, we discuss the results of the estimation. We start by discussing the results for the data from the surveys related to households/individuals, then turn to the results of the estimation performed on the surveys from businesses and conclude by putting the insights from both together.

3.3.1 Technologies for Households and Individuals

The density estimates of the parameters linked to the technologies (for the technologies adopted by households/individuals, i.e. Health Information on the Internet, Internet, Internet Banking, News and Magazines, P2P eCommerce, Private Emails, VoIP/Teleconferencing, Social Networks) can be seen in Figure 1, some of the corresponding numerical values are also listed in Table A.1 in Appendix A (the estimates of the parameters linked to the countries are available in Figure 2 and Table A.2).

The estimates for the speed of adoption (β_j) inherent/specific to the technology can be found in panel (a). Those estimates appear to be consistent with the estimates reported in previous uses of diffusion models. The literature review in Mahajan et al. [1990], for example, reports typical results of diffusion coefficients for durable goods around 0.3 and 0.7, which are however highly contingent on the technology or population being analysed. Here, the coefficients for the fastest diffusing technologies, the internet and internet banking, are estimated to be in that range, while the estimates for the other technologies are lower than that. None of the technologies though appears to be an outlier in terms of the speed of diffusion, which might be expected for a set of technologies that could be considered relatively homogeneous, at least in terms of their main characteristic.

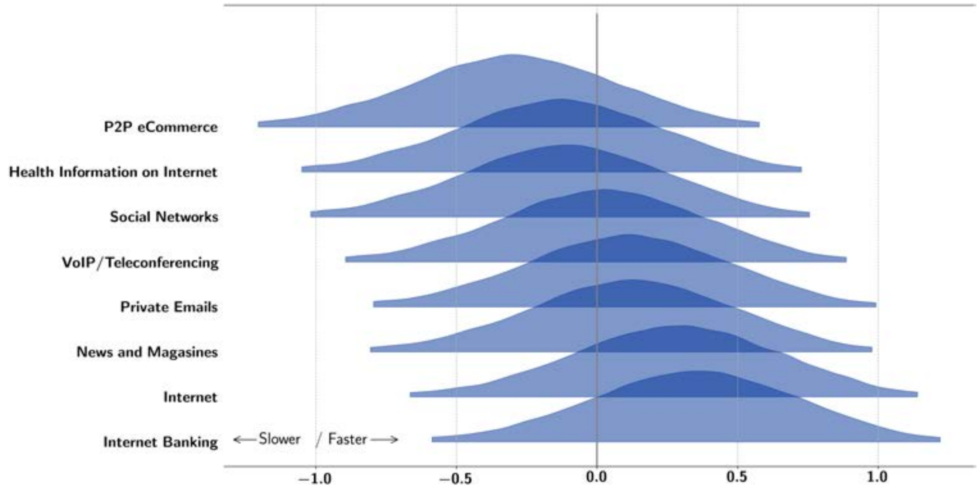
It is striking, however, that there is an outlier for the timing of diffusion (τ_j) - P2P eCommerce, estimated to be "newer" (in the sense that it will reach the half-point of its full diffusion later) than the other technologies, and an outlier also for the maximum adoption (κ_j) - VoIP/Teleconferencing, estimated to have a greater adoption. The estimates for the timing in panel (b) of Figure 1 suggest that P2P eCommerce is newer and diffusing more slowly, while the Internet and Internet Banking, and to some extent also Private Emails, are older and diffusing faster. This seems sensible, given the somewhat greater complexity and specificity of P2P eCommerce relative to those other, less complex and more general, technologies.

The estimates for the maximum adoption (κ_j in panel (c) of Figure 1) signal a particularly large adoption of VoIP/Teleconferencing once they will be entirely diffused. While this may not be surprising per se, it might raise some eyebrows as it implies a higher adoption of VoIP/Teleconferencing (which relies on the internet) than the internet itself. That result might be due to the recent acceleration of adoption in the context of the CoViD pandemic that have the model "confused" in that it considers the technology to still be in the accelerating phase of the diffusion process (one can see in Table 1 that this is indeed the technology for which the progression inside the countries was the largest).[6]

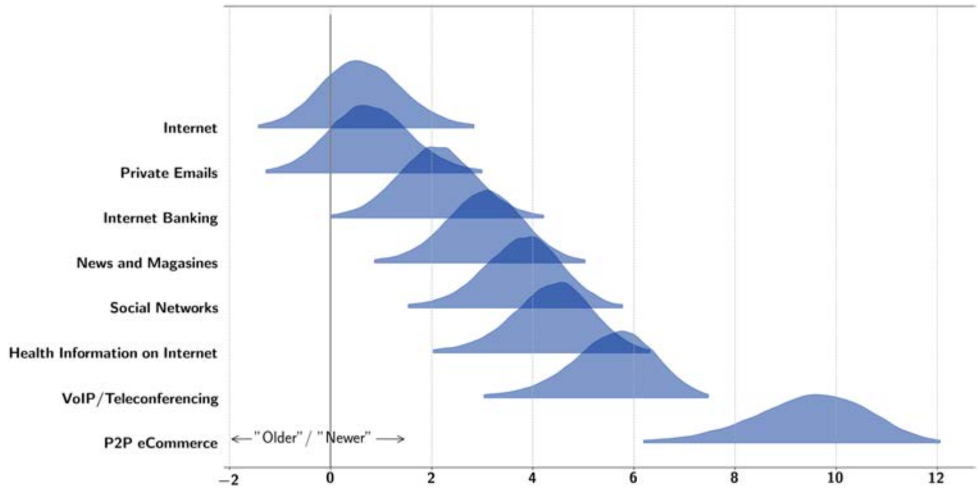
[6] This result also suggests that it is important to take into account the effects of major shocks such as the dot-com bubble, the great financial crisis, the sovereign debt crisis, and the COVID-19 pandemic when monitoring and forecasting digital and, more general, technology diffusion. This will be done in a future paper.

Figure 1: Density estimates of parameters linked to the technologies
(for technologies adopted by households/individuals)

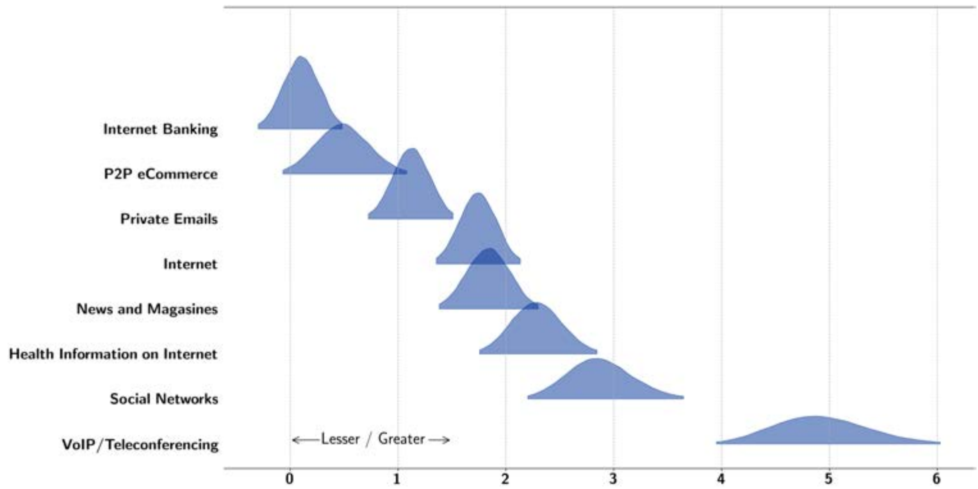
(a) Speed of adoption (β_j)



(b) Timing of adoption (τ_j)



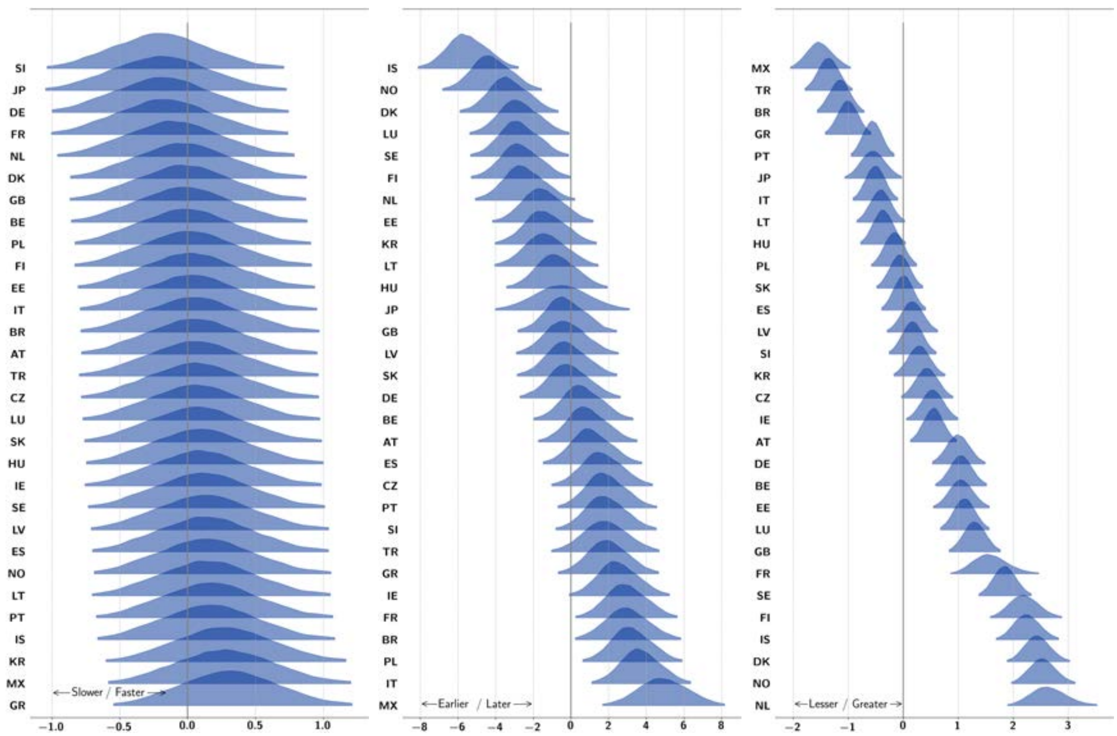
(c) Maximum adoption (κ_j)



The densities estimated for the coefficients related to the countries are shown in Figure 2 as well as Table A.2 in Appendix A (still for technologies adopted by households/individuals). The densities for the speed modifier (δ_i in panel (a)) capture to what extent the country plays a role in explaining a speed of diffusion different from the cross-country average. The estimates are relatively similar across the countries, and the ordering is consistent with the notion that technologies diffuse faster in smaller economies, owing for example to greater population density and/or openness.[7]

Figure 2: Density estimates of parameters linked to the countries
(for technologies adopted by households/individuals)

(a) Speed modifier (δ_i) (b) Timing modifier (θ_i) (c) Maximum modifier (χ_i)



[7]We invite the reader accustomed to frequentist models to analyse those results with caution and keeping in mind that the inference mechanism is very different here. Since represented here are the parameter with their 98% highest density interval, having those interval crossing the $x = 0$ axis does not mean we "reject" a country effect.

The estimates for the timing modifier (θ_i in panel (b) of Figure 2) suggest that adoption begins earlier in countries that are smaller and/or more north, which would broadly match the pattern of location of some of the major technology companies. According to the estimates for the maximum diffusion at the country level (κ_j in panel (c) of Figure 2), the smaller and/or more north countries also tend to have a higher maximum adoption once the technology has fully diffused.

3.3.2 Technologies for Businesses

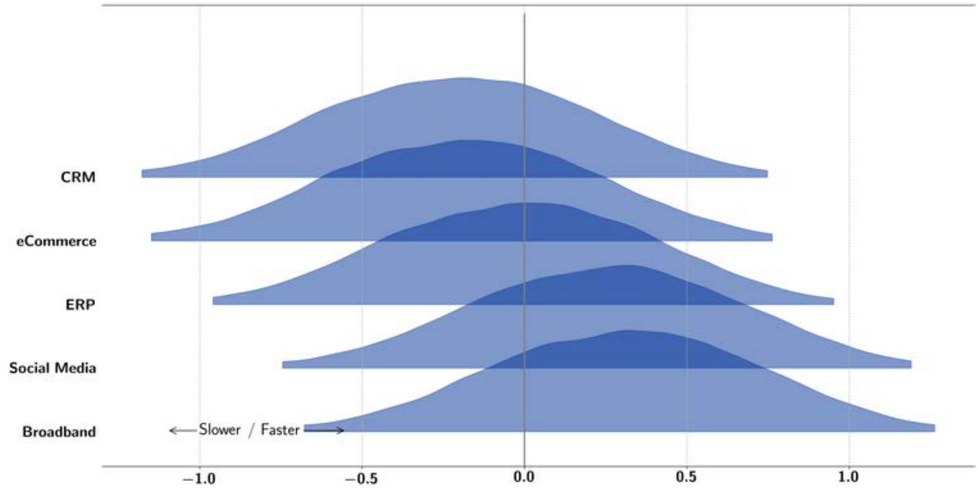
The results on the technologies adopted by businesses (i.e. Broadband, CRM, eCommerce, ERP and Social Media) are available in Figures 3 and 4, as well as Tables A.3 and A.4 in Appendix A. Much like for the technologies used by individuals/households, the speed of diffusion (β_j in panel (a) of Figure 3) seems to be roughly the same for all technologies, and the discriminant factors for technologies appear, once again, to lie in the timing and maximum adoption.

Examining the timing of adoption, it seems that CRM, ERP and the use of Social Media by businesses are expected to reach the half-point of their diffusion somewhere around 2011 - 5 years after the start of the dataset for business technologies, which corresponds to the value 5 of τ_j in panel (b) of Figure 3. Broadband adoption and eCommerce, on the other hand, seem to be technologies that are going to reach the middle point of their adoption at a later date, with eCommerce seemingly reaching the inflexion point of its diffusion curve only around 2020.

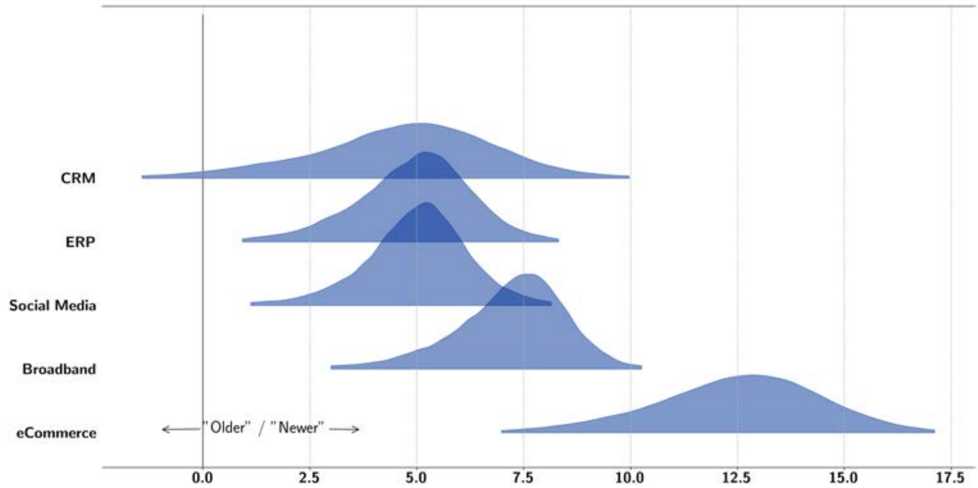
The most striking contrast, however, seems to be found on the side of the maximum adoption (κ_j in panel (c), Figure 3 vs Figure 1). CRM and ERP adoption will likely peak at a lower adoption rate than the other technologies, that are most likely to be widely spread once their diffusion is complete. This makes intuitive sense: CRM and ERP are mostly designed for companies active in industrial and commercial activities. As such, it seems relatively logical that they will not be seen as useful to as many companies as other technologies such as, for instance Broadband internet, and so not diffuse as widely. It seems, therefore, that, in this respect, the Bayesian perspective is helpful to discriminate accurately between technologies.

Figure 3: Density estimates of parameters linked to the technologies
(for technologies adopted by businesses)

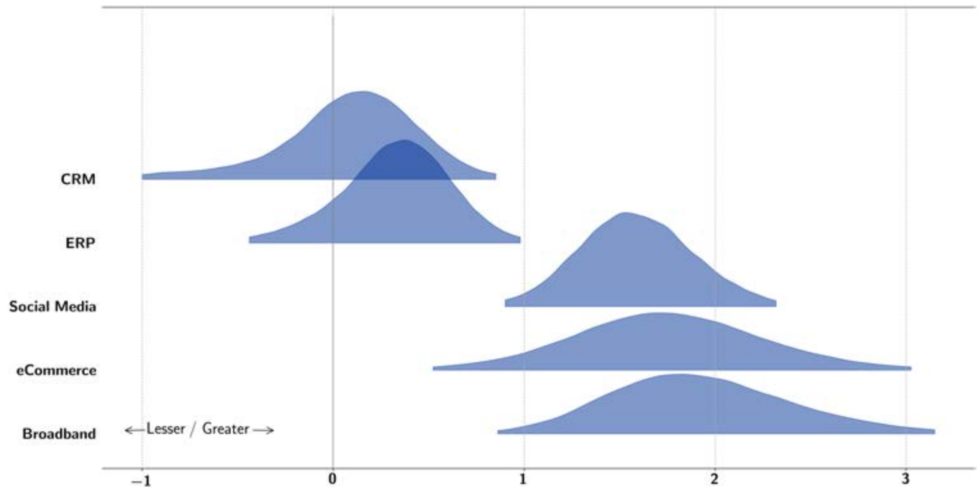
(a) Speed of adoption (β_j)



(b) Timing of adoption (τ_j)



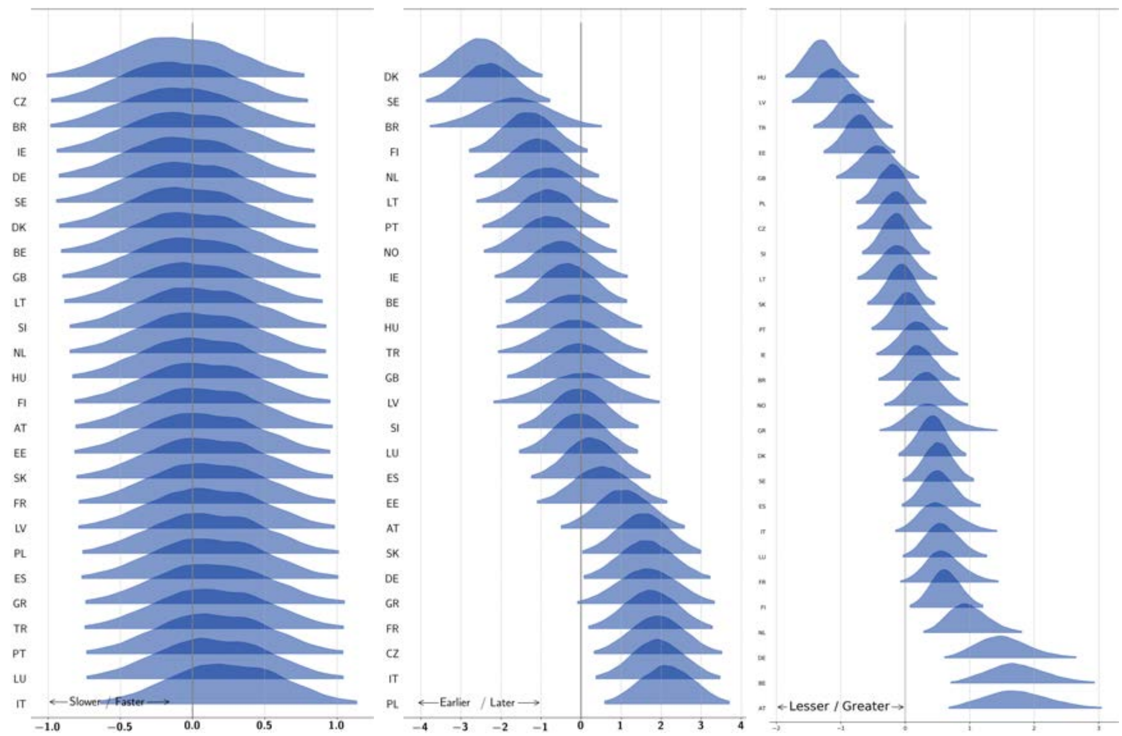
(c) Maximum adoption (κ_j)



Both the spread between countries and the uncertainty on the parameters is larger on the business than on the individuals data set while, if one abstracts from the outliers in each of the data set, the spread of the timing parameter modifier θ_i is comparable. The ranking in terms of means is different from the one for the individuals in its details but it remains broadly similar.

Figure 4: Density estimates of parameters linked to the countries
(for technologies adopted by businesses)

(a) Speed modifier (δ_i) (b) Timing modifier (θ_i) (c) Maximum modifier (χ_i)



4 Forecasts

4.1 Production

In order to produce forecasts using the framework presented above, we use Stochastic Variational Inference (SVI) rather than the more traditional Monte Carlo Markov Chain (MCMC) that was used to produce the estimates in Subsection 3.3. SVI works by approximating the distribution of the latent variables by a multivariate normal distribution and then looking for the set of parameters minimizing the Evidence Lower BOund (ELBO) using stochastic gradient descent. This is done using the Numpyro Python library (Phan et al. [2019], Jha and Saha [2018]).

The choice of using SVI rather than MCMC was made for a practical reason. Stochastic Variational Inference is much faster than Monte Carlo Markov Chain techniques and produces approximately the same results. The drawback is that, contrary to MCMC, we have no guarantee of asymptotic convergence to the true posterior distribution. Since the task of evaluating forecasts requires a large number of estimation steps and that the main purpose of this section is not to discuss the value of the coefficients but rather to analyse the divergence of the forecasted density with the observed value, the choice was made to use this approximate technique.

4.2 Evaluation

The evaluation of the density forecast is done using the Probability Integral Transform (referred to as PIT henceforth). This method, described at length for instance in Diebold et al. [1998], consists of gathering the observed values on the cumulative distribution function (CDF) of the density forecast. If the density forecast were perfect, the density of this collection of values would correspond to the density observed in the case of a uniform distribution (and this would mean that the empirical CDF corresponds to the forecast one). We therefore plot this empirical collection of transformed values against several draws of the same number of uniform random variables. If the empirical curve is above those draws, it means that there are more values that fall on this part of the CDF than what would be normally expected.

Formally, for each observed datapoint for technology j and country i at time t, a_{ijt}, we produce the density forecast $\hat{p}_{ijt|..t-h}$ for the h-steps ahead forecast of technology j in country i, then the probability integral transform set for the h-steps ahead forecasts are:

$$PIT_h = \{F_{ijt|..t-h}(a_{ijt}) \forall i, j, t\}$$

where $F_{ijt|..t-h}$ is the empirical cumulative distribution function:

$$F_{ijt|..t-h}(a_{ijt}) = \int_0^{a_{ijt}} \hat{p}_{ijt|..t-h}(x)dx$$

Since, when a predictive density coincides with the empirical one, their cumulative density functions are identical, the distribution of their results should follow a uniform distribution. We prefer this to summarising this fit through a single numerical statistic such as the statistic of Kullback and Leibler [1951], because kernelised graphical representations of the PIT against samples of the uniform distribution enable making diagnostics about the properties of the predictive density that are causing the divergence.

We now turn to the actual forecasts. We begin with "in-sample" technologies - technologies that can be used to train the model. We do this because the use of the model for forecasting would naturally be envisaged primarily for the technologies that helped to estimate the model, and because it is a setting in which the forecast would be expected to be very accurate. We then consider technologies that cannot be used to train the model. This is important for practical applications in which the model would be of most interest. Indeed, it might be used in the forecasting of the most recent, or completely new, technologies - technologies that would naturally be "out-of-sample" for a few years (before data about their adoption starts to be collected). This is of course a more challenging test for the model, and also a test against overfitting.

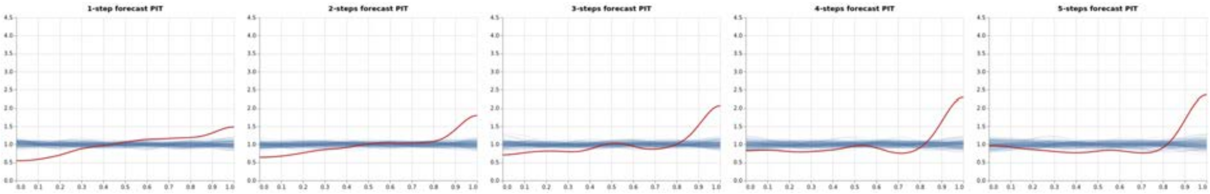
4.3 In-sample technologies

We start by producing the forecast of the next time-steps of technologies that were used to calibrate the model. We do so by "masking" the last iterations of each of the series to emulate the situation where a forecaster wants to produce a forecast of the adoption of technology she already has data on for the next few years.

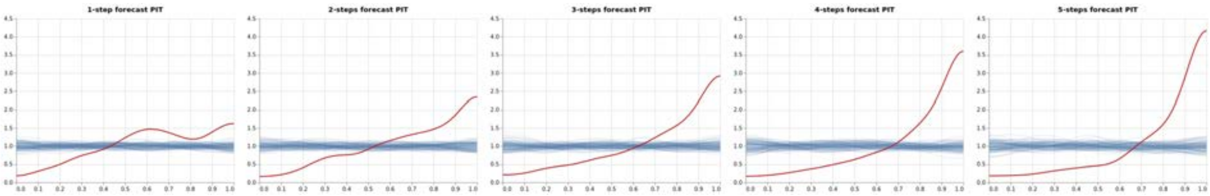
Concretely, and for each of the lags considered here (from $h \in [1, 5]$), we produce $\hat{p}_{ijt|..t-h}$, the h-steps ahead density forecast (where j is the index of the technology and i the index of the country) and their corresponding cumulative density function $\hat{F}_{ijt|..t-h}$. We then compute the value of the CDF for all observed adoption proportion, $\hat{F}_{ijt|..t-h}(a_{ijt})$. This statistic should be indistinguishable from a draw from the Uniform distribution on the $[0, 1]$ interval if the forecast was perfect. Figure 5 shows the statistic against a draw from the Uniform distribution with the same number of samples as the one available as observed data (to allow comparability).

Figure 5: Precision of forecasts for in-sample technologies

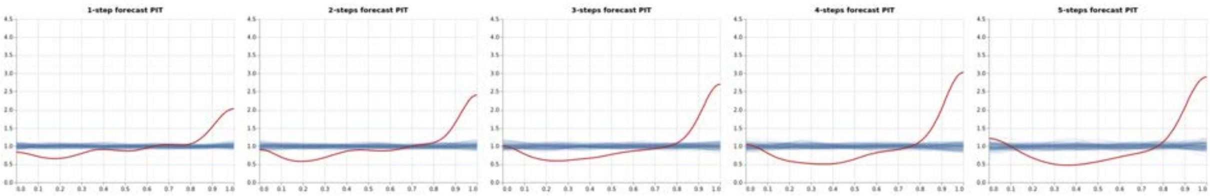
(a) Technologies adopted by households/individuals



(b) Technologies adopted by businesses



(c) All technologies



Notes: The panels show the predictive probability integral transform, PIT, for 1 to 5 steps ahead (red line) relative to a family of 100 for a random sample of identical size from a uniform distribution.

A couple of points on the results in Figure 5 are worth making. First, the results are better for shorter horizons, for which the predictive PITs are more closely aligned to the PITs from a random sample, and worse for longer horizons. Second, when using technologies either directed only at individual consumers or at corporate consumers, in panels (a) and (b), the forecast seems biased: there is a spike of observations for values that are near the maximum of the forecast distribution.

This seems to indicate that the forecast is too conservative: indeed, the growth of the actual adoption leads our generalised logistic function-based guide. It is interesting to note though that when both sets of technologies are added to the model, as shown in panel (c) of Figure 5, the larger volume of data helps providing a better forecasting density. While the size of the bias is reduced substantially in this way, it does not, however, manage to get rid of the bias towards lower values completely. Subsection 4.4 below provides some insights as to the reason for such a result.

4.4 Out-of-sample technologies

Part of the value of the framework developed in this paper lies in its ability to predict adoption of technologies that were not used to train it, and it not being limited to single technologies. Due to the nature of the evolution of technology, once enough information is available at the level of a single technology to produce a valuable known-technology forecast, it may be too late for that forecast to be of much practical or policy relevance. As such, the performance of the model on technologies that are unknown is probably the single most interesting aspect of the forecasting application of this model.

We therefore analyse the one-step-ahead forecast produced by holding the technology out of the training set. We then include the information available for the first year where we have data available on the technology (for any country) and produce the one-step-ahead forecast again, and then progressively add new years to the training set. We repeat the process for 5 years, holding each time a technology out and keeping only the forecast for the one-step-ahead forecast for this technology.

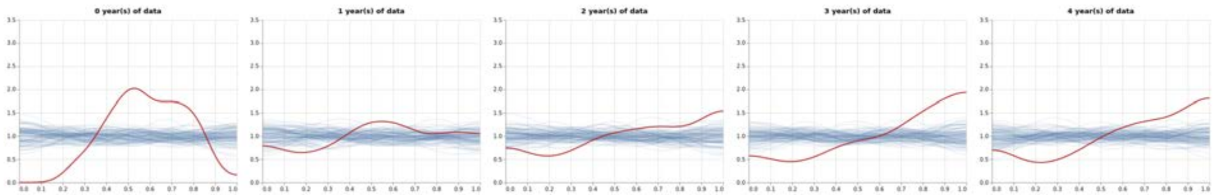
The results, in Figure 6, are sensible, although they seem surprising at first sight. Contrary to the results presented in Subsection 4.3, only the results for the one-step-ahead forecast for out-of-sample technologies show that the model is under-confident in the case of technologies for consumers when it has no information, in the sense that most observed data fall at the center of the predicted distribution (and the PIT is therefore hump-shaped), and overestimating rather than underestimating for technologies destined to producers. With hindsight, this makes sense if the initial data point helps to anchor the forecast.

What is most remarkable, though, is that the framework performs well when we have very little data on a specific technology at our disposal. Indeed, the results for the predictions with one to four years of data available are quite close to the uniform distribution that would be the goal for such a statistic. It seems, therefore, that the framework performs best when applied to recent technologies and is applied to the short-term prediction task.

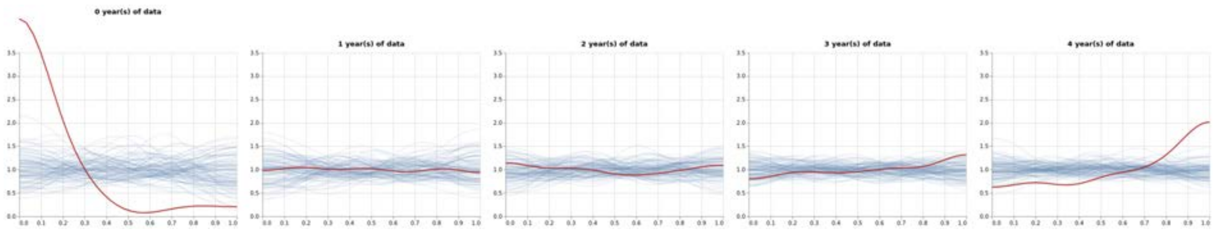
Finally, one may remark that, as more data becomes available, the Probability Integral Transform plot starts to exhibit the bias shown in Subsection 4.3. This might suggest that the regularisation imposed by the priors is a bit too stringent and that looser priors on the speed of adoption might be required. This has to be weighed against the need for priors making the inference task possible, given the dataset we have. Further work is required to study this

Figure 6: Precision of forecasts for out-of-sample technologies

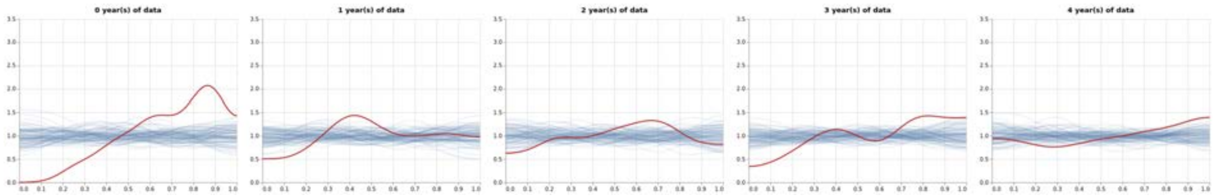
(a) Technologies adopted by households/individuals



(b) Technologies adopted by businesses



(c) All technologies



Notes: The panels show the predictive probability integral transform, PIT, for 1 to 5 steps ahead (red line) relative to a family of 100 PITs for a random sample of identical size from a uniform distribution.

possibility and, perhaps, to provide more flexibility in the design of diffusion curve. Another possibility is that the adoption process is impacted by exogenous shocks.

5 Caveats and Conclusions

This paper has presented a new perspective on technology forecasting, based on a simple yet easily extendable Bayesian hierarchical structure applied to the seminal model of technology diffusion. The key contribution of this paper is thus a methodological one. With the advent of macroeconomic models based increasingly on micro-data, being able to forecast the adoption of individual technologies rather than relying on aggregated features might enable the forecaster to get a more refined understanding of the dynamics of the economy and model it in a more detailed fashion. Further, being able to forecast the diffusion of technologies in certain geographies enables the elaboration of policies with the ultimate goal of facilitating the acquisition, by businesses and individuals alike, of the technology mix that best fits her general strategy.

The paper illustrates that Bayesian methods are well-suited for such a task. Indeed, they can be used even with relatively scarce data and produce predictions that are both easy to understand and allow for the quantification of the uncertainty of the same predictions (given the model). The paper also contributes to the policy debate though, owing to the conclusions from the application to digitalisation, and the data set for EU and OECD countries and different digital technologies adopted by either households/individuals or by businesses.

The first conclusion is that digital technology diffusion speed remains largely dependent on the intrinsic characteristics of said technologies. However, country effects exists, both in the timing of the adoption and also in its speed. The relative ranking of countries might seem unsurprising: more north countries tend to lead, both in terms of speed as in terms of timing, while countries in some other regions tend to adopt technologies later and more slowly.

This article highlights, however, that there are significant differences in terms of the diffusion of technologies directed at individuals compared to the ones directed at businesses. While there seem to be countries in which both businesses and individuals adopt technologies faster and earlier than in other, some countries seem to display disparities in their rank on speed and timing of diffusion between the two sets of technologies. This finding, while requiring further investigation, might be seen as an argument for tackling digitalisation separately in what concerns the supply and the demand side of the economy.

The distinction made in this article between digital technologies adopted by households/individuals or businesses is of course artificial, as many digital technologies are used by both of them, especially so-called "enabling technologies" which provide stepping stones for other, more advanced technologies that would not exist without them. We therefore classified technologies according to the intensity with which the technologies are used by respectively by households and businesses, or the likelihood that they are adopted because they are such stepping stones, as for example broadband as business technology (enabling intelligent process automation, etc).

The framework used in this paper has many attractive features such as a tolerance to missing values inherent to the type of data used here and, up to a point, a certain resistance to outliers (being a Bayesian model, it tends to discount them and not let them influence the rest of the inference as much as other techniques might). Moreover, it is a framework making it possible to produce forecasts even for technologies not previously seen, something that is intended to be looked at in related work on-going. The initial forecasts become better as more technologies are introduced to the model and could potentially be enhanced with the addition of regressors about the characteristics of the technologies.

While that framework has many advantages, it also has a number of features that could be viewed as shortcomings, some inherent to the methods used, and some that may be lifted in future. Among the first, linked to the Bayesian perspective, the necessity of choosing priors is the most obvious. Although the authors took care to define uninformative priors were relevant (mostly on the variance parameters), such priors depend inevitably on both technical and domain knowledge. In the context of this analysis, several priors where analysed with essentially the same results (sometimes at the cost of longer computing time). The interested reader is, of course, welcome to estimate the model with a different set of priors and verify that the conclusions essentially remain.

Among the limitations that could be lifted in future endeavours are the issues revealed by the PITs, notably at the margins of the densities. Indeed, determining whether they come from outliers that would warrant a separate modelling or from structural properties of the beta likelihood function could make the model perform better not least at prediction tasks. Moreover, additional insights into the deeper characteristics of the technologies would make the task of applying the framework to other, notably new technologies, easier. In general, the literature on technological forecasting, because it often attempts to predict diffusion of technologies in isolation (either by virtue of the dataset or by the model having no dependence of coefficients across technologies), and has not taken into account the intrinsic characteristics of the geographical areas in which they are diffusing.

The framework resulting from the new perspective on technology diffusion is, in any case, already usable in practice to make predictions on the near term evolution of digitalisation in different geographical areas and, more generally, to enable inference on digital diffusion. The inference is also the next focus in this line of research notably in relation to potentially disruptive episodes or events that might impact on technology diffusion, such as the financial and sovereign debt crises, the COVID-19 pandemic or more recent ones.

References

Claudio Baccianti, Vincent Labhard, and Jonne Lehtimäki. Digitalisation, institutions and governance, and diffusion: Mechanisms and evidence. *ECB Working Paper*, (2675), 2022.

Eli Bingham, Jonathan P. Chen, Martin Jankowiak, Fritz Obermeyer, Neeraj Pradhan, Theofanis Karaletsos, Rohit Singh, Paul Szerlip, Paul Horsfall, and Noah D. Goodman. Pyro: Deep Universal Probabilistic Programming. *arXiv preprint arXiv:1810.09538*, 2018.

Diego Comin and Bart Hobijn. An exploration of technology diffusion. *American Economic Review*, 100(5):2031–2059, December 2010. doi: 10.1257/aer.100.5.2031. URL https://www.aeaweb.org/articles?id=10.1257/aer.100.5.2031.

Francis X. Diebold, Todd A. Gunther, and Anthony S. Tay. Evaluating density forecasts. *International Economic Review*, 39(4):863–883, November 1998. URL https://www.sas.upenn.edu/~fdiebold/papers/paper16/paper16.pdf.

Silvia Ferrari and Francisco Cribari-Neto. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815, 2004. doi: 10.1080/0266476042000214501. URL https://doi.org/10.1080/0266476042000214501.

A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 2013. ISBN 9781439840955. URL https://books.google.de/books?id=ZXL6AQAAQBAJ.

Zvi Griliches. Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 25(4):501–522, 1957. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1905380.

Matthew D. Hoffman and Andrew Gelman. The no-u-turn sampler: Adaptively setting path lengths in Hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(1):1593–1623, January 2014. ISSN 1532-4435.

Ashutosh Jha and Debashis Saha. Diffusion and forecast of mobile service generations in Germany, UK, France and Italy - a comparative analysis based on Bass, Gompertz and simple logistic growth models. In Peter M. Bednar, Ulrich Frank, and Karlheinz Kautz, editors, *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018, Portsmouth, UK, June 23-28, 2018*, page 17, 2018. URL https://aisel.aisnet.org/ecis2018_rp/17.

John K Kruschke. Bayesian data analysis. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(5):658–676, 2010.

Solomon Kullback and Richard A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951. doi: 10.1214/aoms/1177729694. URL https://doi.org/10.1214/aoms/1177729694.

Jonathan Lee, Peter Boatwright, and Wagner A Kamakura. A Bayesian model for prelaunch sales forecasting of recorded music. *Management Science*, 49(2):179–196, 2003.

Peter J. Lenk and Ambar G. Rao. New models from old: Forecasting product adoption by hierarchical Bayes procedures. *Marketing Science*, 9(1):42–53, 1990. doi: 10.1287/mksc.9.1.42. URL https://doi.org/10.1287/mksc.9.1.42.

Vijay Mahajan, Eitan Muller, and Frank M. Bass. New product diffusion models in marketing: A review and directions for research. *Journal of Marketing*, 54(1):1–26, 1990. doi: 10.1177/002224299005400101. URL https://doi.org/10.1177/002224299005400101.

Edwin Mansfield. Technical change and the rate of imitation. *Econometrica*, 29(4):741–766, 1961. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1911817.

Du Phan, Neeraj Pradhan, and Martin Jankowiak. Composable effects for flexible and accelerated probabilistic programming in NumPyro. *arXiv preprint arXiv:1912.11554*, 2019.

FJ Richards. A flexible growth function for empirical use. *Journal of Experimental Botany*, 10 (2):290–301, 1959.

Paul Stoneman. *Soft Innovation: Economics, Product Aesthetics, and the Creative Industries*. Number 9780199697021 in OUP Catalogue. Oxford University Press, 2011. ISBN ARRAY(0x4cf76d08). URL https://ideas.repec.org/b/oxp/obooks/9780199697021.html.

Tatcha Sudtasan and Hitoshi Mitomo. Comparison of diffusion models for forecasting the growth of broadband markets in Thailand, 2017. URL http://hdl.handle.net/10419/168541.

H. Peyton Young. Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American Economic Review*, 99(5):1899–1924, December 2009. doi: 10.1257/aer.99.5.1899. URL https://www.aeaweb.org/articles?id=10.1257/aer.99.5.1899.

Appendix A: Tables

Table A.1: Parameter estimates for (intrinsic) speed (β) and timing (τ) of adoption (for technologies adopted by households/individuals)

	beta					tau				
	mn	std	med	5%	95%	mn	std	med	5%	95%
Technology										
P2P eCommerce	-0.31	0.38	-0.30	-0.94	0.31	9.42	1.26	9.53	7.25	11.24
Health Information	-0.14	0.38	-0.14	-0.78	0.47	4.34	0.90	4.40	2.84	5.66
Social Networks	-0.13	0.38	-0.12	-0.76	0.48	3.76	0.88	3.81	2.31	5.09
VoIP/Teleconferencing	0.01	0.38	0.02	-0.62	0.62	5.52	0.95	5.61	3.90	6.88
Private Emails	0.10	0.38	0.11	-0.53	0.71	0.77	0.88	0.75	-0.59	2.23
News and Magasines	0.11	0.38	0.12	-0.52	0.72	3.01	0.86	3.04	1.62	4.33
Internet	0.27	0.38	0.28	-0.37	0.88	0.62	0.89	0.59	-0.74	2.08
Internet Banking	0.34	0.38	0.35	-0.30	0.95	2.09	0.86	2.09	0.74	3.45

Note: 'P2P' stands for 'peer-to-peer', 'VoIP' for 'voice-over-IP', 'Health Information' for 'Health Information on the Internet', 'mn' stands for 'mean', 'std' for 'standard deviation', 'med' for 'median'.

Table A.2: Parameter estimates for (country-specific) speed (δ) and timing (θ) modifier (for technologies adopted by households/individuals)

	delta					theta				
	mn	std	med	5%	95%	mn	std	med	5%	95%
Country										
SI	-0.19	0.37	-0.20	-0.79	0.42	1.12	0.60	1.10	0.15	2.12
JP	-0.18	0.37	-0.18	-0.77	0.44	-0.08	0.60	-0.09	-1.04	0.93
DE	-0.18	0.38	-0.19	-0.79	0.45	-0.31	0.87	-0.31	-1.74	1.13
FR	-0.16	0.37	-0.17	-0.76	0.46	1.81	0.61	1.79	0.82	2.82
NL	-0.10	0.37	-0.10	-0.69	0.52	-1.56	0.60	-1.58	-2.51	-0.55
DK	-0.03	0.37	-0.03	-0.62	0.59	-2.06	0.60	-2.08	-3.00	-1.05
GB	-0.02	0.37	-0.02	-0.61	0.60	-0.19	0.58	-0.22	-1.12	0.78
BE	-0.02	0.37	-0.02	-0.61	0.60	0.36	0.59	0.34	-0.57	1.34
PL	-0.00	0.37	-0.01	-0.60	0.61	1.97	0.60	1.96	1.01	2.96
FI	0.02	0.37	0.02	-0.57	0.64	-0.94	0.61	-0.96	-1.91	0.07
EE	0.02	0.37	0.01	-0.58	0.64	-1.66	0.60	-1.68	-2.62	-0.65
IT	0.03	0.37	0.03	-0.56	0.65	2.27	0.60	2.26	1.30	3.27
BR	0.05	0.37	0.05	-0.55	0.67	1.82	0.65	1.81	0.76	2.90
AT	0.06	0.37	0.05	-0.54	0.67	0.51	0.58	0.49	-0.43	1.48
TR	0.06	0.38	0.06	-0.54	0.69	1.12	0.67	1.12	0.04	2.24
CZ	0.07	0.37	0.06	-0.52	0.69	-1.72	0.60	-1.74	-2.68	-0.71
LU	0.07	0.37	0.06	-0.53	0.68	1.00	0.60	0.99	0.03	2.01
SK	0.08	0.37	0.07	-0.52	0.69	-0.15	0.60	-0.17	-1.11	0.85
HU	0.10	0.37	0.09	-0.50	0.72	1.53	0.60	1.51	0.57	2.53
IE	0.10	0.37	0.09	-0.50	0.72	-0.48	0.60	-0.50	-1.45	0.52
SE	0.12	0.37	0.12	-0.47	0.74	-1.72	0.59	-1.74	-2.65	-0.73
LV	0.13	0.37	0.13	-0.47	0.75	-0.19	0.61	-0.21	-1.17	0.83
ES	0.14	0.37	0.14	-0.45	0.76	0.64	0.58	0.63	-0.29	1.62
NO	0.14	0.37	0.14	-0.45	0.76	-2.62	0.61	-2.65	-3.59	-1.60
LT	0.15	0.37	0.15	-0.44	0.77	-0.83	0.63	-0.85	-1.85	0.22
PT	0.17	0.37	0.17	-0.42	0.79	1.12	0.58	1.10	0.18	2.09
IS	0.18	0.37	0.17	-0.42	0.79	-3.41	0.63	-3.43	-4.42	-2.34
KR	0.28	0.38	0.27	-0.33	0.90	-0.85	0.61	-0.88	-1.84	0.17
MX	0.28	0.38	0.28	-0.33	0.92	3.02	0.77	3.01	1.78	4.30
GR	0.32	0.37	0.31	-0.28	0.94	1.20	0.61	1.19	0.22	2.21

Note: The countries are denoted by their two-digit ISO codes, 'mn' stands for 'mean', 'std' for 'standard deviation', 'med' for 'median'.

Table A.3: Parameter estimates for (intrinsic) speed (β) and timing (τ) of adoption (for technologies adopted by businesses)

	beta					tau				
	mn	std	med	5%	95%	mn	std	med	5%	9%
Technology										
CRM	-0.22	0.42	-0.21	-0.91	0.46	4.68	2.31	4.83	0.63	8.13
eCommerce	-0.20	0.42	-0.19	-0.88	0.48	12.45	2.11	12.61	8.80	15.58
ERP	-0.02	0.42	-0.01	-0.70	0.65	4.89	1.49	5.03	2.35	7.05
Social Media	0.24	0.42	0.25	-0.45	0.92	4.97	1.36	5.07	2.69	6.98
Broadband	0.30	0.42	0.31	-0.40	0.97	7.13	1.46	7.32	4.55	9.11

Note: 'CRM' stands for 'Customer Relationship Management', 'ERP' for 'Enterprise Resource Planning', 'mn' stands for 'mean', 'std' for standard deviation, 'med' for 'median'.

Table A.4: Parameter estimates for (country-specific) speed (δ) and timing (θ) modifier (for technologies adopted by businesses)

	delta					theta				
	mn	std	med	5%	95%	mn	std	med	5%	95%
Country										
NO	-0.10	0.39	-0.11	-0.73	0.54	1.90	0.69	1.91	0.77	3.03
CZ	-0.10	0.39	-0.11	-0.73	0.54	-0.80	0.72	-0.80	-1.97	0.39
BR	-0.08	0.40	-0.09	-0.73	0.58	-1.65	0.91	-1.65	-3.15	-0.17
IE	-0.06	0.39	-0.07	-0.69	0.58	-0.50	0.72	-0.50	-1.67	0.69
DE	-0.05	0.39	-0.05	-0.68	0.59	-2.31	0.67	-2.31	-3.41	-1.19
SE	-0.05	0.39	-0.06	-0.68	0.58	1.62	0.69	1.62	0.49	2.76
DK	-0.04	0.39	-0.05	-0.67	0.59	-2.51	0.67	-2.51	-3.60	-1.41
BE	-0.02	0.39	-0.03	-0.65	0.62	-0.35	0.66	-0.35	-1.44	0.72
GB	-0.01	0.39	-0.02	-0.64	0.63	-0.08	0.77	-0.08	-1.35	1.19
LT	0.00	0.39	-0.01	-0.63	0.64	-0.84	0.76	-0.84	-2.09	0.40
SI	0.04	0.39	0.03	-0.59	0.68	-1.10	0.68	-1.10	-2.21	0.02
NL	0.04	0.39	0.03	-0.59	0.67	-0.06	0.65	-0.06	-1.13	1.00
HU	0.05	0.39	0.04	-0.58	0.69	-0.24	0.79	-0.24	-1.55	1.03
FI	0.06	0.39	0.05	-0.57	0.70	-1.30	0.64	-1.30	-2.36	-0.24
AT	0.06	0.39	0.06	-0.57	0.70	1.06	0.67	1.05	-0.05	2.16
EE	0.07	0.39	0.06	-0.56	0.70	0.53	0.71	0.53	-0.62	1.69
SK	0.09	0.39	0.08	-0.54	0.73	1.54	0.65	1.55	0.47	2.62
FR	0.10	0.39	0.10	-0.53	0.74	1.75	0.68	1.74	0.63	2.86
LV	0.11	0.39	0.10	-0.52	0.75	-0.06	0.89	-0.04	-1.56	1.37
PL	0.12	0.39	0.12	-0.50	0.76	0.26	0.64	0.25	-0.80	1.32
ES	0.12	0.39	0.11	-0.51	0.75	2.14	0.68	2.14	1.01	3.25
GR	0.14	0.39	0.13	-0.50	0.79	1.65	0.75	1.65	0.42	2.88
TR	0.14	0.39	0.14	-0.49	0.79	-0.19	0.80	-0.17	-1.53	1.11
PT	0.15	0.39	0.15	-0.47	0.79	-0.04	0.65	-0.04	-1.11	1.03
LU	0.15	0.39	0.14	-0.48	0.79	-0.84	0.69	-0.84	-1.98	0.29
IT	0.24	0.39	0.24	-0.39	0.88	1.91	0.67	1.91	0.80	3.02

Note: The countries are denoted by their two-digit ISO codes, 'mn' stands for 'mean', 'std' for 'standard deviation', 'med' for 'median'.

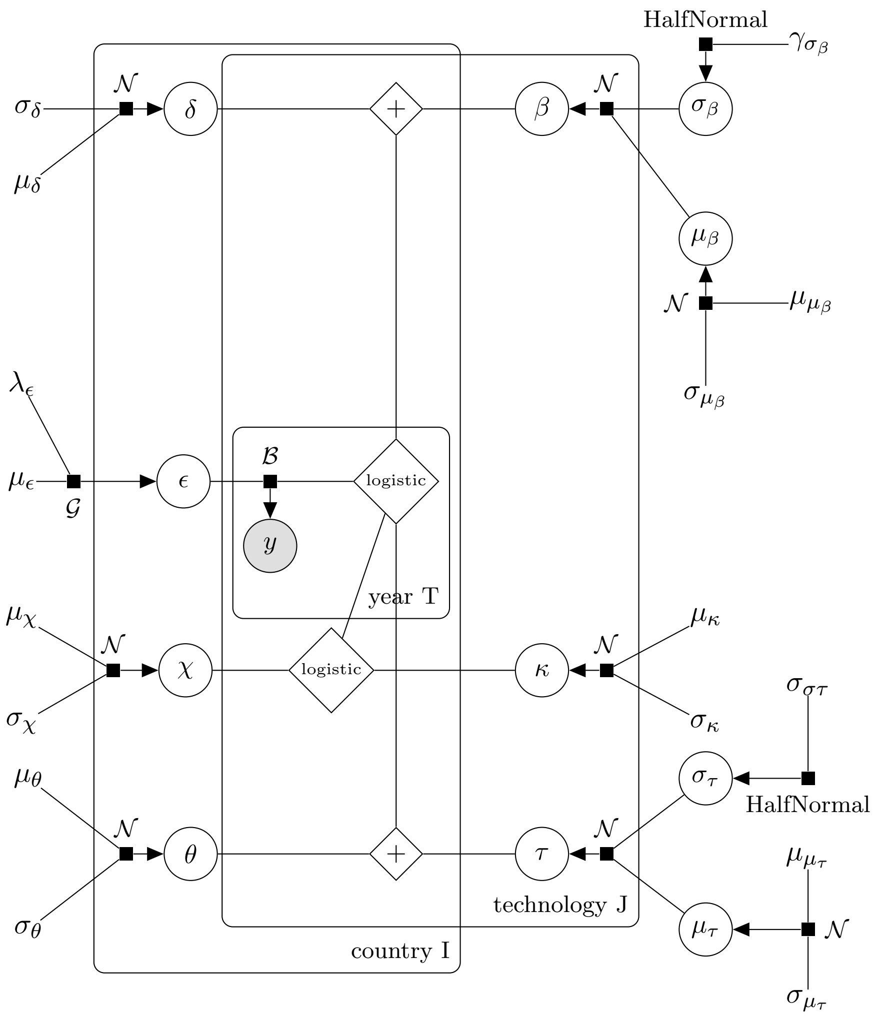
Table A.5: The priors in the model

	Distribution
Coefficient	
Concentration (error) (ϵ_i)	Gamma(5.0, 1.0)
Shape (ω_j)	Uniform(0.5, 1.5)
Maximum Adoption (χ_i and κ_j)	Normal(0.0, 1.0)
Speed of diffusion (δ_i and β_j)	Normal(hierarchical priors)
Timing of diffusion (θ_i and τ_j)	Normal(hierarchical priors)
Hierarchical Priors	
Hierachical mean parameters	Normal(0.0, 1.0)*
Hierarchical variance parameters	HalfNormal(5.0)

Note: *Except for the the prior on τ_j which was set to a Normal(0.0, 10.0) to allow for the model to fit in the middle of the dataset.

Appendix B: Figures

Figure B.1: Complete directed factor model



Acknowledgements

The work has benefitted from comments by P. Lenk (University of Michigan), and by colleagues and/or seminar participants at the European Central Bank and the Université libre de Bruxelles. In particular, we would like to thank R. Anderton. The code for replicating the analysis in this paper is available on request from the authors.
The views expressed in this paper are those of the authors and do not necessarily reflect those of any of the institutions the authors are or were affiliated to.

Charles Hoffreumon
Université libre de Bruxelles, Brussels, Belgium; email: charles.hoffreumon@ulb.ac.be

Vincent Labhard
European Central Bank, Frankfurt am Main, Germany; email: vincent.labhard@ecb.europa.eu

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-5285-9 ISSN 1725-2806 doi:10.2866/632240 QB-AR-22-065-EN-N