université PARIS-SACLAY

ENSAE ParisTech

Two essays on the market for bitcoin mining and one essay on the fixed effects logit model with panel data

NNT : 2018SACLG002

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'école nationale de la statistique et de l'administration économique

École doctorale n°578 Sciences de l'Homme et de la Société (ED SHS)
Spécialité de doctorat: Sciences Économiques

Thèse présentée et soutenue à Paris, le 4 septembre 2018, par

Benjamin Walter

Composition du Jury :

Cyril Grunspan Responsable majeure ingénierie financière, ESILV	Président
Winfried Koeniger Professor of economics, Université de Saint-Gall	Rapporteur
Jean-Marc Robin Professeur d'économie, Sciences Po	Rapporteur
Bruno Biais Directeur de recherche, CNRS, HEC	Examinateur
Xavier D'Haultfœuille Administrateur hors classe, CREST, ENSAE	Directeur de thèse
Julien Prat Professeur d'économie, CNRS, École Polytechnique	Directeur de thèse

Thèse de doctorat

Contents

Résumé substantiel en français **4**

Remerciements **6**

General introduction **8**

1 An equilibrium model of the market for bitcoin mining **12**

1.1 Introduction 12

1.2 Bitcoin and the market for mining 16

1.3 The Model 19

1.4 Calibration 25

1.5 Extensions 35

1.5.1 Model with halvings. 35

1.5.2 Mothballing and scrapping options. 38

1.6 Conclusion 46

1.7 Appendix 47

2 How to make the Bitcoin network more environmentally friendly **54**

2.1 Introduction 54

2.2 Bitcoin and the miners 56

2.3 The model 59

2.4 Numerical analysis 67

2.5 Conclusion 75

3 Identification and estimation of the average marginal effect in a panel data fixed effects logit model **76**

3.1 Introduction . . . 76
3.2 Identification results . . . 78
3.2.1 The parameters of interest . . . 78
3.2.2 Sharp bounds . . . 79
3.2.3 A nicer characterization for the bounds . . . 83
3.2.4 Computation of the bounds . . . 87
3.3 Estimation . . . 89
3.3.1 The use of an index . . . 89
3.3.2 A first idea . . . 90
3.3.3 A second idea . . . 91
3.3.4 More than two periods are available . . . 92
3.4 Monte-Carlo simulations . . . 92
3.5 Conclusion . . . 95

General conclusion **96**

Bibliography **102**

Résumé substantiel en français

Ma thèse se compose de deux parties indépendantes. La première, qui comporte deux chapitres, appartient au nouveau champ dit de "la crypto-économie", tandis que la deuxième, qui constitue le troisième chapitre, s'occupe d'économétrie théorique.

Dans le premier chapitre, co-écrit avec M. Julien Prat, nous présentons un modèle qui prédit la puissance de calcul totale déployée par les mineurs de bitcoins en utilisant le taux de change bitcoin / dollar. Il s'agit du premier modèle dynamique d'équilibre du marché des mineurs. Le problème auquel les mineurs font face s'apparente à un quasi-cas d'école d'un problème d'investissement dans un secteur où les revenus sont incertains. En effet, le marché des mineurs remplit quatre conditions précises. Premièrement, de par la nature des problèmes cryptographiques que doivent résoudre les mineurs, ce secteur bénéficie de rendements d'échelle constants. Deuxièmement, miner étant une activité à la portée de tous, il y a libre entrée sur ce marché. Troisièmement, les mineurs ont tous accès à la même technologie et font tous face aux mêmes coûts. Si leur revenus, qui dépendent directement du taux de change bitcoin / dollar, sont très volatils, ceux-ci constituent la seule source d'incertitude. Il n'y a aucun aléa au niveau individuel. Enfin, le matériel informatique utilisé étant spécifique à cette activité, il ne peut être revendu. Acheter une de ces machines constitue donc un investissement irréversible. Ces quatre caractéristiques correspondent aux hypothèses du modèle de Caballero et Pyndick (1996). Il y a cependant une différence notable entre ce dernier et le marché des mineurs: le matériel de minage bénéficie d'un important taux de progrès technique. Nous adaptons donc le modèle précédent en conséquence et le calibrons avec les données. Nous démontrons empiriquement sa pertinence en comparant ses prédictions avec la puissance de calcul observée.

Dans le deuxième chapitre, écrit seul, je propose une méthode simple et réaliste pour enrayer la hausse (à court terme) et amorcer la baisse (à plus long terme) de la consommation déraisonnable d'électricité de Bitcoin et rompre le lien entre celle-ci et le taux de change du bitcoin. En m'appuyant sur une version simplifiée du modèle développé dans le premier chapitre, je mets en évidence l'inefficacité du protocole Bitcoin actuel. Je montre que celui-ci offre un niveau de sécurité beaucoup trop élevé au prix d'un formidable gâchis d'électricité financé par les détenteurs de bitcoins. En effet, il serait possible de diminuer grandement la puissance de calcul du réseau tout en conservant à un niveau extrêmement faible la probabilité qu'un agent

malveillant puisse réussir à annuler une transaction. Cela signifie qu'il serait possible de moins rémunérer les mineurs en leur accordant moins de nouveaux bitcoins pour chaque nouveau bloc trouvé et ainsi converger plus rapidement vers le protocole de long terme proposé par Nakamoto (2008), selon lequel les mineurs sont exclusivement rémunérés par les frais de transaction. Sans même tenir compte de l'externalité environnementale, les détenteurs de bitcoins bénéficieraient d'une telle mesure car celle-ci réduirait l'augmentation du stock total de bitcoins. Un tel protocole Pareto-dominerait le protocole actuel car la libre entrée sur le marché des mineurs assure que ces deniers ne réalisent pas de profit, quelque soit le protocole en vigueur. Malheureusement, la transition entre les deux protocoles s'avère impossible à cause d'un problème d'incitations. Les mineurs déjà entrés sur le marché ne sont, eux, pas indifférents entre les protocoles car ils ont déjà acheté leurs machines. Ce sont justement eux qui décident si le changement de protocole doit avoir lieu ou non. Pour pallier ce problème, je détermine le protocole le plus économe en électricité qui ne lèse pas les mineurs actifs. L'intuition est la suivante: afin que ceux-ci acceptent une future et permanente baisse de leurs revenus, il faut leur accorder le plus tôt possible une prime conséquente. Ainsi, il faut commencer par augmenter fortement et temporairement le nombre de nouveaux bitcoins créés par blocs avant de lui donner sa valeur souhaitée pour le plus long terme.

Dans le troisième chapitre, j'étudie les problèmes d'identification et d'estimation de l'effet marginal moyen dans un modèle logit sur données de panel avec effets fixes. Je commence par montrer que le paramètre d'intérêt n'est que partiellement identifié et je pose le problème d'optimisation qu'il faudrait résoudre pour obtenir les bornes exactes de l'intervalle d'identification. Il s'agit d'un problème d'optimisation sur un espace fonctionnel, extrêmement complexe en théorie comme en pratique. En m'appuyant sur la théorie des systèmes de Chebyshev, je montre que les bornes admettent une caractérisation simple qui permet de les calculer très rapidement numériquement. Malheureusement, le problème dit "du fléau de la dimension" empêche, la plupart du temps, d'appliquer la même méthode pour estimer les bornes sur des données. Je propose alors plusieurs idées pour contourner ce problème et les évalue à l'aide de simulations.

Remerciements

En priorité, je tiens à remercier du fond du cœur mes deux codirecteurs de thèse, Xavier D'Haultfœuille et Julien Prat. Julien m'a apporté une aide précieuse dans le moment le plus difficile de ma thèse en m'encadrant sur le nouveau sujet sur lequel j'ai choisi de travailler. Il m'a beaucoup appris en matière de rigueur et de clarté dans l'écriture d'un article scientifique. C'est grâce à cette persévérance de sa part que le premier chapitre de ma thèse a tant de panache. Enfin, il m'a toujours accordé beaucoup de temps et a fait de nombreux efforts pour me faire rencontrer les sommités de mon domaine. De son côté, Xavier s'est montré extrêmement disponible au début de ma thèse en interrompant toujours immédiatement la tâche qu'il avait commencée pour répondre à toutes mes questions. Il a joué un rôle fondamental dans l'élaboration du dernier chapitre de ma thèse et m'a montré comment donner un caractère transdisciplinaire à la recherche en allant dénicher un résultat de mathématiques pures méconnu pour l'appliquer à notre problème.

Je souhaite remercier tout particulièrement Laurent Davezies, pour sa constante bonne humeur et pour le temps quasi infini qu'il a passé à répondre à mes questions les plus farfelues.

Je voudrais aussi remercier Daniel Augot pour sa bienveillance, pour l'intérêt porté à mes travaux ainsi que pour ses efforts dans le but de créer des projets communs rassemblant les différents acteurs de la Blockchain.

Je remercie également Josef Zweimüller et Sven Seuken pour leur accueil chaleureux à Zurich et mon intégration facilitée dans leurs équipes respectives.

Je remercie le CREST d'avoir financé ma thèse. Je garde un excellent souvenir de ces trois années et de la bonne ambiance qui règne dans le bâtiment. Une telle atmosphère est avant tout le fruit des professeurs qui ne se prennent pas au sérieux, à commencer par Francis Kramarz, et qui instaurent un rapport d'égalité avec les doctorants, contrairement à ce qui se pratique généralement dans les autres universités. Je tiens ainsi à remercier Christophe Bellego, Pierre Cahuc, Edouard Challe, Bruno Crépon, Laurent Davezies, Xavier D'Haultfœuille, Bertrand Garbinti, Robert Gary-Bobo, Aneth John, Thierry Kamionka, Olivier Loisel, Jacques Mairesse, Jacques Melitz, Manasa Patnam, Julien Prat, Anna Simoni, Arne Uhlendorff, Thibaud Vergé et Michael Visser. Bien sûr, je souhaite également remercier les autres doctorants

du CREST pour les moments inoubliables passés ensemble. Un grand merci à Ivan Ouss, qui a dû me supporter tous les jours pendant deux ans et puis à mes autres collègues de bureaux, à savoir Morgane Cure, Hugo Molina et Julien Monardo, puis Sandro Favre, Christian Kiedaisch, Lei Li, Tobias Renkin et Daphné Skandalis. Merci aussi à Reda Aboutajdine, Hélène Benghalem, Antoie Bertheau, Marianne Blehaut, Arthur Cazaubiel, Antoire Ferey, Sebastian Franco Bedoya, Lucas Girard, Aymeric Guidoux, Morgane Guignard, Malka Guillot, Yannick Guyonvarch, Jeremy Hervelin,
Melina Hillion, Thomas Jagelka, Alexis Larousse, Clémence Lenoir, Jeremy L'Hour, Alicia Marguerie, Esther Mbih, Andreea Minea, Sandra Nevoux, Louis Pape, Bérangère Patault, Fabien Perez, Julie Pernaudet, Audrey Rain, Anasuya Raj, Emilie Sartre, Clémence Tricaud, Jérôme Trinh, Antoine Vatat, Ao Wang et Jiekai Zhang.

General introduction

Specific case for this dissertation This dissertation can be seen as a specific case since it concatenates two independent topics. In fact, I started working on a theoretical econometrics subject but, due to great difficulties, then switched to another topic which fascinates me: Bitcoin. Such a choice was motivated by the novelty and the originality of my new topic, which raises incredibly many unanswered questions in all the sub-fields of economics. Today, I reckon I made the right decision since this subject has become extremely fashionable and working on it has enabled me to write my dissertation on time. To deliver a faithful image of the dissertation, the introduction splits in two parts: one part on Bitcoin and the other one on econometrics.

Bitcoin

Bitcoin was created in the beginning of January 2009. It is the first currency which works without any central actor such that a central bank or commercial banks. But Bitcoin is more than a mere currency. Each payment is assorted with a condition that the recipient must fulfill in order to redeem the sent funds. This technology, known as "smart contracts", can be used to solve many commitment problems. For a long time, Bitcoin remained extremely confidential or confined to a small circle of computer science experts. That is why, until 2016, when Bitcoin eventually became more famous for the general public, most works on it were undertaken by computer sciantists. The aim of those works was often to improve Bitcoin's functionalities. For instance, King and Nadal (2012) show how to deprive Bitcoin from its dependency on energy consumption. Assia et al. (2013) create a protocol which enables users to transfer the ownership of any asset using Bitcoin. Back et al. (2014) design a general method used to leverage Bitcoin's capabilities, which Lerner (2015) resort to in order to enhance the possibilities offered by smart contracts. Finally, Poon and Thaddeus (2016) explain how to transfer bitcoins instantaneously and safely. Other articles focus on the safety of transactions. Karame et al. (2012) wonder to which extent it is possible to cancel a Bitcoin payment, when the corresponding good is delivered immediately. Reid and Harrigan (2012) question the anonimity of transactions. Decker and Wattenhofer (2013) explain how delays when transmitting information on the network can be detrimental and Grunspan and Pérez-Marco (2017) correct Nakamoto's computations about the probability of success of an attack.

Along with Bitcoin's democratization, its disruptive potential and the one of blockchain, the underlying technology, have become ever more palpable. Little by little, one could witness the development of a whole ecosystem of more than a thousand cryptocurrencies endowed with different aims and functionnalities, and applications relying on those cryptocurrencies. New economic practices have appeared and attracted economists' attention who have started studying the unidentified economic object with the seemingly most natural tool: monetary economics. Fernández-Villaverde and Sanches (2016) wonder whether it makes sense to have a plethora of cryptocurrencies. Hong et al. (2017) study to what extent cryptocurrencies can represent a threat for fiat currencies. Chiu and Koeppl (2017) call into question the choice for parameters of the Bitcoin protocol and Gandal et al. (2017) shed light on some exchange rate manipulations.

True, Bitcoin enables the development of new economic practices. But its link to economics is much stronger. It is fascinating to see to which extent its design hinges on the economic science. If Bitcoin's safety remains, before all, a matter of robustness of its cryptographic primitives, the security also crucially depends on miners' economic incentives to provide the network with their computing power. Several articles study those incentives. Rosenfeld (2011) study the different mining pools reward systems, Biais et al. (2017) wonder under which condition can the chain of blocks split in two and Huberman et al. (2017) focus on the behavior of the different actors once (almost) all the bitcoins are mined.

Those last articles are probably the closest to my work since I focus on miners' behavior as well. But what especially stroke me was Bitcoin's staggering electricity consumption, which, in June 2018, was a bit above 0.3% of the world electricity consumption. In a first chapter, I try to understand how miners make their decisions about when investing in new mining hardware. Since the ₿/ $ exchange rate directly influences miners' revenues, the investment decision must obviously depend on it. Yet, knowing exactly when miners should invest remains more challenging because they must take into account the high volatility of the exchange rate, the high embodied technical progress rate the mining hardware enjoys and the competitive nature of the market. In the first chapter, my coauthor, Mr. Julien Prat, and I design a model which accounts for those points. Using the ₿/ $ exchange rate, the model manages to reproduce fairly well the observed computing power on the medium and long terms.

In a second chapter, I start with describing the current Bitcoin protocol to show how much electricity and money are wasted for the sake of an unnecessarily high level

of security. I present the ideal protocol but show that it remains utopic due to an impossible transition from the current one. Indeed, any protocol change must be agreed upon by miners. Miners would certainly veto this specific proposal since they would be the only losers. Finally, I introduce the best realistic protocol: the one which leads to the smallest amount of electricity being consumed without hurting miners.

Panel data with a binary dependent variable

As opposed to statistics, econometrics often deals with endogeneity issues. Resorting to panel data remains one of the most frequent strategy, all the more so since panel data are ever more available in many areas. In econometric models, the use of panel data enables the econometrician to decompose the unobserved component, or error, in an individual-specific variable, fixed in time and an idiosyncratic random variable. If only the time-fixed component[1], called "fixed effect", suffers from endogeneity, then estimators remain consistent as long as there exists a way to transform the model so as to make the fixed effects disappear. Note that getting rid of the fixed effects is a necessity. Simply estimating them does not work because of the so-called "incidental parameters problem". Lancaster (2000) provides a literature review on this issue, which arises when the number of parameters to estimate grows with the sample size, at the same speed. In general, it implies the non-consistency of the estimator of the parameter of interest.

With the linear model, the most frequently used, one can easily get rid of the fixed effects simply by differentiating the data with respect to time. In this chapter, I consider the binary dependent variable case. The marginal effect of an explanatory variable on the dependent variable is interpreted as the extent to which the explanatory variable affect the probability that the dependent one be equal to one. When this probability is already very close to zero or very close to one, the marginal effect of an explanatory variable has to be extremely week. Thus, for the model to make mathematical and economic sense, the marginal effect of a variable must tend to zero at infinity. The linear model, which, in this case, remains the simplest option, does not satisfy this basic requirement.

[1]The term "panel data" is used when several observation are available for each individual. If hese multiple observations often stem from a survey that the same people answer several times, panel models can be used in many other situations. For the sake of simplicity, I shall always speak about time.

With panel data, one must prefer the logit model since it satisfies the above requirement and still enables the econometrician to get rid of the fixed effects, resorting to a simple sufficient statistic. Chamberlain (2010) shows that when only two periods are available, only the logit model fulfills these two conditions.

Yet, for many applications, knowing the value of the parameter of the model is not enough since it cannot easily be interpreted. Here, I focus on the average marginal effect of an explanatory variable, which is, according to me, one of the most interesting quantities. The average marginal effect cannot be directly computed with the model parameter. Indeed, since individuals marginal effects depend on the values of those individuals' fixed effects, the average marginal effect depends on the joint distribution of the fixed effects and the explanatory variables... which is not identified. The easiest solution is then to assume that the fixed effects follow a normal distribution are independent from the explanatory variables. But this is a very restrictive assumption.

Here, I try to remain as agnostic as possible on the joint distribution of the fixed effects and the explanatory variables. Such a choice comes at a cost: the average marginal effect is no longer point identified but only set identified. Finding the sharp identification interval boils down to minimize and maximize the average marginal effect on the set of all the fixed effects distributions which satisfy the constraints imposed by the model. But except for some very particular cases, infinite dimension optimization problems, like the one we have here, are extremely difficult to solve in practice. Chernozhukov et al. (2013) suggest a solution which involves an arbitrary discretization the set of distributions. Of course, their method under-estimates the real size of the identification region.

My two coauthors for this chapter, Messrs. Davezies and D'Haultfœuille, and I manage to solve this problem. We provide a simple, easy-to-use and non computer-intensive method to identify the sharp bounds of the identification region of the average marginal effect. This is the main contribution of our article.

1 An equilibrium model of the market for bitcoin mining

Joint work with Julien Prat.

Abstract

We propose a model which uses the Bitcoin/US dollar exchange rate to predict the computing power of Bitcoin's network. We show that free entry places an upper-bound on mining revenues and we devise a structural framework to measure its value. Calibrating the model's parameters allows us to accurately forecast the evolution of the network computing power over time. We establish the accuracy of the model through out-of-sample tests and investigation of the entry rule. We find that around one third of seigniorage income is dissipated in electricity consumption. The model indicates that a slowing down in the rate of technological progress will significantly increase Bitcoin's carbon footprint.

1.1 Introduction

Bitcoin is the first currency that operates without a central authority or a trusted third party. It enables merchants and customers to transact at a low cost and almost as securely as if they were relying on the banking system. The disintermediation of monetary transactions is only the initial stage of the paradigm shift initiated by Bitcoin. Its success has ushered in a new era of financial innovation, with hundreds of cryptocurrencies created over the last couple of years.

Bitcoin's design relies on a hybrid model that combines the robustness of its cryptographic primitives with the economic incentives of the agents participating in the execution of its protocol. Miners are at the center of the infrastructure since they guarantee the validity of transactions. They stack transactions into blocks and timestamp those in a cryptographically robust way by adding a "proof-of-work".[2] Miners are rewarded for their efforts with new bitcoins and transaction fees. The cost of attacking Bitcoin is proportional to the computing power deployed by miners because it determines the difficulty of the cryptographic puzzles included in their proofs-of-work.

As the value of Bitcoin skyrocketed, so did the resources devoted to mining. What started as a hobby for a few miners using their personal computers, eventually blos-

[2] See Section 1.2 for a description of the tasks accomplished by miners.

somed into an industry which consumes nearly 0.3% of the world's electricity through its network of mining farms,[3] each one of them operating thousands of machines specially designed for mining. In spite of the growing concerns about the carbon footprint of the mining industry, our paper is the first to propose an equilibrium model characterizing its evolution over time. We show that miners' investment in computing power can be accurately forecasted using the Bitcoin/US dollar (₿/$) exchange rate.

Investment in mining hardware has two important characteristics. First, it cannot easily be reversed: machines have no resale value outside of the market for mining because they have been optimized for mining only. Second, there is a lot of uncertainty about future revenues due to the tremendous volatility of ₿/$ exchange rate. This combination generates a range of inaction where expected revenues are too low to justify entry, yet still sufficient to prevent incumbents from exiting the market.

The main challenge for our analysis is that we cannot consider the problem of each miner in isolation or treat revenues as exogenous. Instead, we have to take into account how returns are endogenously determined by the number of active miners. A key insight of our model is that Bitcoin's protocol generates revenues functions that are decreasing in aggregate output, thereby ensuring that the market for mining behaves as a competitive industry.

Combining ₿/$ exchange rate with the total computing power of Bitcoin's network, we construct a new measure for miners' payoffs. Our model predicts that miners buy hardware only when this measure reaches a reflecting barrier. It never exceeds the barrier because new entries push down payoffs by triggering additional increases in mining costs. The characterization of the equilibrium is complicated by the fact that mining hardware benefits from a high rate of embodied technological progress. We show how one can adapt the canonical model of Caballero and Pyndick (1996) to account for this trend, and prove that the entry barrier decays at the rate of technological progress.

Then we calibrate the model and find that it forecasts remarkably well how miners respond to changes in ₿/$ exchange rate. The accuracy of our predictions is a testament to the fact that miners operate in a market where perfect competition is a good approximation of reality. Its structure verifies many properties that are often assumed but rarely verified in practice. First, free entry holds because mining is a mostly un-

[3] See, among other sources, digiconomist.net/bitcoin-energy-consumption .

regulated activity with a streamlined set of tasks. To enter the mining race, one simply has to buy the appropriate hardware and install the mining software. Second, there is very little heterogeneity among miners since they all face the same problem and earn the same rewards. Third, as explained below, the mining technology exhibits returns to scale that are constant by nature. Fourth, the elasticity of revenues with respect to the network computing power is commonly known because it is encoded in Bitcoin's protocol, and is therefore observable by all parties. Finally, we have access to perfectly clean and exhaustive data since all transactions are public. The conjunction of all these features is extremely rare, if not unique, thus making the market for Bitcoin mining a perfect laboratory for models of industry dynamics.

After having established that our baseline model accurately matches the data, we relax a couple of simplifying assumptions in order to assess its robustness. First, we allow for discontinuities in miners' rewards that take into account reductions in the monetary creation rate triggered by Bitcoin's protocol every four years. Second, instead of assuming that investment is completely irreversible, we endow miners with the option to mothball or scrap their machines. We find that these extensions improve the fit of the model but only marginally so. However, the model with partially reversible investment has the significant benefit of separating the entry from the operating costs, indicating that around one third of seigniorage income is dissipated in electricity consumption. Studying the impact that each parameter has on the electricity consumption of miners, we find that Bitcoin's carbon footprint is likely to increase, principally because of a slowdown in the rate of progress of the mining technology.

Related literature.—Our paper uses insights from the literature on irreversible investment to contribute to the nascent field of *cryptoeconomics*. Bitcoin was created almost a decade ago when Nakamoto's paper (Nakamoto (2008)) was made public on October 31st 2008. It did not immediately attract much attention and it took a few years for Bitcoin to become the focus of academic research. Early works analyzed the reliability of Bitcoin's network (Karame et al., 2012; Decker and Wattenhoffer, 2013). Reid and Harrigan (2012) examined the anonymity of users, which enabled Athey et al. (2017) to quantify the different ways bitcoins are used and Foley et al. (2018) to precisely identify illegal bitcoin users. Grunspan and Pérez-Marco (2017) and Bowden et al. (2018) both corrected mathematical approximations made by Nakamoto in his seminal paper.

It is only recently that papers studying the economic implications of cryptocur-

rencies have started to emerge. Most articles rely on monetary economics for their analysis. Observing the plethora of existing cryptocurrencies, Fernández-Villaverde and Sanches (2016) wonder under which conditions competition between currencies is economically efficient and how those currencies should be regulated. Hong et al. (2017) and Schilling and Uhlig (2018) study the interactions between fiat and crypto currencies. Chiu and Koeppl (2017) assess the choice of values for the parameters that underlie Bitcoin's design while Gandal et al. (2017) analyze exchange rate manipulations. Cong and He (2017) question the public disclosures of information which result from the use of the blockchain technology.

A series of recent papers is more closely related to our research since it studies the market for mining. Rosenfeld (2011), Houy (2016) and Biais et al. (2017) investigate miners' incentives to behave cooperatively, as expected in Bitcoin's protocol, or to play "selfish". Ma et al. (2018) model the market for mining as a game between miners. Huberman et al. (2017) look at the very long run, when miners will be rewarded in transaction fees only. They analyze how users set their fees and how their decisions impact electricity consumption.

Our paper also models the market for mining but unlike aforementioned articles, we focus on miners' entry decisions. We show that their behavior can be captured using standard methods from the real options literature. Since it would be impossible to cover all the major contributions to this field, we refer to Dixit and Pyndick (1994) and their bibliography for a broad overview. Our model being devised in an equilibrium setting, it builds on the seminal work of Bertola and Caballero (1994) and Caballero and Pyndick (1996) on industry dynamics. We find that, despite its apparent novelty, the market for Bitcoin mining behaves very much like a competitive industry. Our analysis illustrates that it is a perfect laboratory for real options theory because all the miners solve a common problem whose parameters are easily observable.

Structure of the paper.—The article is organized as follows. Section 1.2 briefly explains how the market for Bitcoin mining operates. Section 1.3 introduces our baseline model that yields the computing power of the network as a function of ₿/$ exchange rate. Section 1.4 presents the data and explains how to calibrate the model. Section 1.5 proposes two extensions of the baseline model that relax its simplifying assumptions. Section 1.6 concludes. The proofs of the Propositions and some additional results are relegated to the Appendix.

1.2 Bitcoin and the market for mining

This section describes the tasks accomplished by miners and the rewards they get in return. Since it is well beyond the scope of this paper to explain the overall architecture of Bitcoin, we only cover the elements that are required for the understanding of our model, and refer readers interested in a more comprehensive treatment to Nakamoto (2008) and Antonopoulos (2014).

The function of miners.— Bitcoin is a decentralized cryptocurrency which operates without a central authority. Decentralization is achieved through the recording of transactions in a public ledger called the blockchain. The main challenge for a decentralized currency is to maintain consensus among all participants on the state of the blockchain (who owns what) in order to prevent double spending of the same coin. A user spends a coin twice when one of her payments is accepted because the recipient is not aware of a previous payment spending the same coin. In order to avoid such conflicts, transactions are added to the blockchain by blocks and producing a valid block is made so difficult that the time it takes to build a block is, on average, much longer than the time it takes for a block to propagate across the network. This ensures that, in most instances, the whole network agrees on which transactions are part of the blockchain.

Blocks are cryptographically chained according to their dates of creation. This incremental process implies that the information contained in a given block cannot be modified without updating all subsequent blocks. Nakamoto's groundbreaking insight was to recognize that the cost of manipulation would increase dramatically in the number of modified blocks, thus ensuring that tampering with a given block becomes prohibitively expensive as more blocks are added on top of it.

To be accepted by other Bitcoin users, a new block must be stamped with a "proof-of-work". Each block possesses a header, which contains both a "nonce", i.e. an arbitrary integer, and a statistic summing up the transactions of the block, the time the block was built and the header of the previous block. Finding a valid proof-of-work boils down to finding a nonce satisfying the condition $h(header) \leq t$, where h is the SHA-256 hash function applied twice in a row and t is a threshold value. The hash function h has the property of being numerically *non-invertible*. Moreover knowing $h(n)$, for any $n \in \mathbb{N}$, yields no information on $h(m)$ for all $m \neq n$. Hence the only way to find a valid nonce is to randomly hash guesses until the condition above is satis-

fied. This activity is called mining and it requires few special skills besides the means to spend resources on the mining process. The average time it takes to mine a valid block can be made arbitrarily long by lowering the threshold t. Since Bitcoin's protocol specifies that one valid block should be found every 10 minutes, the threshold is updated every 2016 blocks to account for changes in the computing power, or hashpower, deployed by miners .

Building a valid block is costly both in terms of hardware and electricity, hence miners must be rewarded. For each block there is a competition between miners. Only the first miner who finds a valid nonce wins the reward: she earns a predetermined amount of new coins and the sum of the mining fees granted by the transactions included in the block. The amount of new bitcoins for block number B is approximately $50 \times (1/2)^{\lfloor B/210{,}000 \rfloor}$ while fees are freely chosen by users.[4] The amount of new coins halves every 210,000 blocks so as to ensure that the supply of bitcoins converges to a finite limit, namely 21 millions.

The market for Bitcoin mining.— To enter the mining race, a potential miner has to buy the right hardware. Free entry prevails because anyone can easily order the machine and install the software. There is no heterogeneity across miners besides their amounts of hashpower and the price they pay for electricity. The amount of hashpower a miner owns should not matter due to constant returns to scale: two pieces of hardware will generate valid blocks exactly twice as often as a single piece of hardware. It has nonetheless become common for miners to build impressive mining farms. Although such concentrations of computing power suggest that returns are increasing,[5] the conclusion is not warranted. It might very well be that the size of each farm is determined by the amount of cheap electricity that is available in its specific location. Such a constraint would determine the geographical allocation of mining power without affecting the industry dynamics at the world level. Actually, since Bitcoin mining is still far from having exploited all the world supply of cheap electricity, it is only natural to conjecture that active miners face operating costs that are broadly similar.

The hardware used for mining benefits from constant upgrades. At the beginning, miners used to mine with their own computers. In mid-2010, they realized that Graph-

[4] We use $\lfloor \cdot \rfloor$ to denote the integer part, i.e. $\lfloor x \rfloor = \max_{n \in \mathbb{N}} \{n \leq x\}$.

[5] Among the reasons why mining may exhibit increasing returns, the most commonly advanced one is that average maintenance costs are decreasing in farm size. Assessing this channel requires detailed micro data on mining which, to the best of our knowledge, are not yet available.

ical Processing Units (GPU) were much more efficient. One year later, miners started using Field Programmable Gate Arrays (FGPA) and, since 2013, they mostly mine with Application Specific Integrated Circuits (ASIC). Investing in a GPU was a reversible decision since GPUs could serve many other purposes besides mining; should the ₿/$ exchange rate drop, the GPU could easily be sold to some video games addict. By contrast, buying an ASIC is an *irreversible investment* because, as indicated by their names, ASICs can be used for Bitcoin mining only. Hence, if the exchange rate collapses, ASICs cannot be resold at a profit because all miners face the same returns.

A mining cycle unfolds as follows. A new miner buys a recent piece of hardware and starts mining with it. Little by little, the revenues generated by her machine drop as ever more powerful hardware enters the race. When the flow income falls below the cost of electricity, the miner turns off her machine and exits the market.

Mining solo is very risky since a miner earns her reward solely when she finds a valid block, which is a very rare event given the number of miners participating in the race.[6] This is why miners pool their resources and share the revenues earned by their pools according to the relative hashpower of each member. Obviously miners have the option but not the obligation to exchange their bitcoins against fiat money. However, since the exchange rate ensures that traders are indifferent between holding fiat money or bitcoins, the value of the reward at the time it is earned is accurately measured by its level in fiat money.[7]

For the sake of completeness, it is worth mentioning that the bitcoins issued with a new block cannot be exchanged straight away. A retention period is imposed because valid blocks are not always added to the blockchain. The validity of the nonce is not enough to maintain consensus when two blocks are found within a short time lapse by two different miners. Then participants will have different views of the state of the blockchain depending on which of the two blocks was broadcasted to them in the first place. Such conflicts create forks in the blockchain that are eventually resolved as miners coordinate on the branch requesting the greatest amount of hashpower ("longest chain rule"). The blocks that were added to the abandoned branch, called "orphan blocks", are discarded. To ensure that the new coins contained in orphan blocks do

[6] On July 1st 2017, the best ASIC miner could perform 14 tera hashes per second and cost $ 2400. The whole Bitcoin network performs 10 exa hashes (10 millions of tera hashes) per second.

[7] Depending on their locations, some investors may convert their bitcoins into different fiat currencies. However, those differences are negligible since from 2009 on, the price of Bitcoin has been far more volatile than that of any major currency.

not contaminate the blockchain, miners have to wait until 100 additional blocks have been added on top of their block before being able to transfer their newly earned coins. In other words, miners have to wait on average 16 hours 40 minutes before transferring their rewards. In practice, this delay is long enough to ensure that the block is indeed included in the blockchain. For our model's purpose, however, forking is a sufficiently rare event that its impact on miners' payoffs can be safely ignored.[8][9]

1.3 The Model

We now propose a framework which captures the main features of the market for mining described in the previous section. Our approach takes the demand for bitcoins as given and uses the trajectory of the exchange rate to predict the hashpower of the network. We devise our model in continuous time and normalize the length of a period to 10 minutes because it corresponds to the average duration separating successive blocks. Since returns to scale are constant, we can think of miners as infinitesimal units of hashpower, and thus assume that the total hashrate of the network takes any positive value on the real line.[10]

Miners' payoffs.—We use R_t to denote the block reward in dollars, i.e. ₿/$ exchange rate multiplied by the sum of new coins and fees. We also let Π_t denote the Poisson rate at which one miner finds a valid block. Then the *flow payoff* P_t of a miner is equal to

$$P_t \equiv R_t \times \Pi_t. \tag{1}$$

The block finding rate Π_t is adjusted every 2016 blocks by Bitcoin's protocol. The updating rule takes the overall hashpower of the network over the previous period as given, and adjusts the difficulty of the hashing problem until new blocks are created on average every ten minutes. This procedure ensures that monetary creation proceeds at a steady pace. Then the complexity of the hashing problem is adjusted on average

[8] Orphan blocks account for less than 0.2% of all mined blocks. The longest chain ever orphaned for a normal reason (not due to a bug) was 4 blocks long, well below 100.

[9] We will also neglect merged mining, i.e. the possibility to mine namecoins together with bitcoins without any additional effort. The reward miners get from namecoins is negligible (not even 0.1%) when compared to the reward in Bitcoins.

[10] Consider, for instance, the following normalization: one miner performs exactly one hash per period, the time interval being 10 minutes. Its relative size is indeed tiny since in mid 2017 the network performed around 10^{19} hashes every second.

every two weeks only. Since our model is designed in continuous time, adding this discrete interval makes it impossible to derive tractable solutions. This is why we slightly idealize the actual protocol and assume that Π_t is continuously adjusted.

Assumption 1. *The valid-proof of work threshold is continuously updated according to the actual total hashrate, so that* $\Pi_t = 1/Q_t$ *for all* t.

The number of hashes the network needs to perform to find a valid block follows a geometric distribution with parameter Π_t. Since the network computes Q_t hashes in one period (10 minutes), the network expected waiting time is $Q_t/\Pi_t = 1$, as prescribed by the protocol. We show in Appendix 1.7 that, during our period of study, the number of blocks mined every day mostly remains within the confidence interval of the null hypothesis. In other words, our data do not significantly deviate from the idealized updating state that would prevail under Assumption 1.

Value of hashpower.—Mining is a costly activity. To operate a unit of hashpower bought at time τ, miners incur the flow electricity cost C_τ. The costs vary with the vintages of the machines because they benefit from embodied technological progress, as newer machines are able to perform more hashes with the same amount of energy.[11] We have already emphasized that investment in hashpower is irreversible because machines cannot be resold. We strengthen this constraint and assume that miners cannot turn off their machines. This simplifying assumption will be relaxed in Section 1.5.2.

Assumption 2. *Mining units cannot be voluntarily switched off so as to save on electricity costs.*

Assumption 2 allows us to express the value of a unit of hashpower of vintage τ as follows

$$V(P_t, \tau) = E_t\left[\int_t^\infty e^{-r(s-t)} P_s ds\right] - \frac{C_\tau}{r}, \tag{2}$$

where r is the discount rate.[12] We have assumed that all the miners of a given vintage face the same problem. In practice, electricity costs may differ across locations but,

[11] Note that we implicitly assume that the price of electricity remains constant. It is easy to relax this restriction by letting C also depend on the current date t. However, changes in electricity costs can be safely ignored in the empirical analysis because they are dwarfed by variations in Bitcoin's exchange rate.

[12] The discount rate r includes the rate at which hardware breaks down. Such failures seem to occur at a much slower rate than technological obsolescence since we do not observe that the network hashpower decreases in the absence of market entry.

due to free entry, only those miners that have access to the cheapest sources of electricity will find it profitable to enter the market. This is consistent with the observation that mining is concentrated in a few places, most notably in China, where electricity is comparatively cheap, or in Nordic countries and Canada, where cold weather makes it easy to cool down the mining rigs.

Under Assumption 1, the flow payoff is given by

$$P_t = R_t/Q_t. \tag{3}$$

Equation (3) defines an isoelastic payoff function with unitary elasticity. Its microfoundation is rather unique since the decreasing relationship between payoffs P and industry output Q does not stem from the satiation of consumers' demand, but is instead generated by the increase in mining costs encoded in Bitcoin's protocol.

We do not attempt to endogenize the demand for bitcoins and thus take the exchange rate R as given. Following much of the literature on irreversible investment, we assume that $(R_t)_{t\geq 0}$ is a Geometric Brownian Motion (GBM hereafter). We assess the accuracy of this assumption when we estimate the model in Section 1.4.[13]

Assumption 3. *$(R_t)_{t\geq 0}$ follows a Geometric Brownian Motion so there is an $\alpha \in \mathbb{R}$, and a $\sigma \in \mathbb{R}_+$, such that*

$$dR_t = R_t\left(\alpha dt + \sigma dZ_t\right), \tag{4}$$

where $(Z_t)_{t\geq 0}$ is a standard Brownian motion.

Knowing the law-of-motion followed by the exchange rate does not enable us to compute the expected value of payoffs because they also depend on the hashpower of the network Q, whose level is endogenously determined. To solve for the equilibrium, one has to simultaneously derive the process followed by Q and the entry policy of miners.

Market entry.—Entrants that want to join the mining race have to buy a unit of hashpower whose price we denote by I_t. Both the entry cost and the cost of electricity decrease over time because machines become more efficient. Let A_t measures technological efficiency, so that a miner can buy A_t units of hashpower at time t with the amount needed to buy one unit of hashpower at date 0. For the reasons explained

[13] Note that the GBM specification disregards the halvings of the money creation rate occurring every 2016 blocks. We will address this shortcoming in Section 1.5.1.

below, we focus on periods where technological improvements accrue at a constant pace, i.e. $A_t = \exp(at)$ with $a > 0$.

Assumption 4. *Machines get more efficient at the constant rate* $a > 0$. *Hence the entry and operating costs satisfy* $I_t = I_0/A_t = \exp(-at)I_0$ *and* $C_t = C_0/A_t = \exp(-at)C_0$.

Free entry ensures that no profits can be made by adding hashpower to the network. Thus the following inequality must hold

$$I_t \geq E_t\left[\int_t^\infty e^{-r(s-t)}P_s ds\right] - \frac{C_t}{r} = V(P_t, t), \text{ for all t.} \tag{5}$$

At times where miners enter the market, (5) will hold with equality. Since the exchange rate follows a Markov process, it is natural to conjecture that their decisions will only depend on the current realization of P: whenever payoffs reach some endogenously determined threshold $\overline{P}_t$, a wave of market entries will ensure that the free entry condition (5) is satisfied.

To see why such a mechanism defines a competitive equilibrium, it is helpful to decompose the law of motion of P. Reinserting (4) into (3) and using Ito's lemma, we find that

$$d\log(P_t) = \left(\alpha - \frac{\sigma^2}{2}\right)dt + \sigma dZ_t - d\log(Q_t). \tag{6}$$

Payoffs are decreasing in Q because the response of the protocol to an increase in total hashpower is to decrease the valid proof-of-work threshold, thus making it less likely for each miner to earn a reward. This is why free entry places an upper bound on payoffs. Their value can never exceed a threshold $\overline{P}_t$ as more miners would find it profitable to enter the market, which would push payoffs further down.

Industry equilibrium.—So far, the main takeaway from our analysis is that the market for mining can be described as a perfectly competitive industry with irreversible investment because Bitcoin's protocol generates a cost function that is increasing in the hashpower of the network. Thus we expect to observe equilibria similar to the ones studied by Caballero and Pyndick (1996) in their seminal paper on industry evolution.

Definition 1.1 (Industry equilibrium)**.**
An industry equilibrium is a payoff process P_t *and an upper barrier* $\overline{P}_t$ *such that:*
(i) $P_t \in [0, \overline{P}_t]$.
(ii) The free entry condition (5) *is satisfied at all points in time, and it holds with equality*

whenever $P_t = \overline{P}_t$.
(iii) The network hashpower Q_t *increases only when* $P_t = \overline{P}_t$.

From a formal standpoint, the only fundamental difference between our model and standard s-S models is that, due to embodied technological progress, entry and variable costs decrease over time. Hence the entry barrier $\overline{P}_t$ cannot remain constant. However, if we impose Assumption 4, so that mining efficiency improves at a constant rate, we can solve for the equilibrium in the space of detrended payoffs in order to recover a flat barrier.

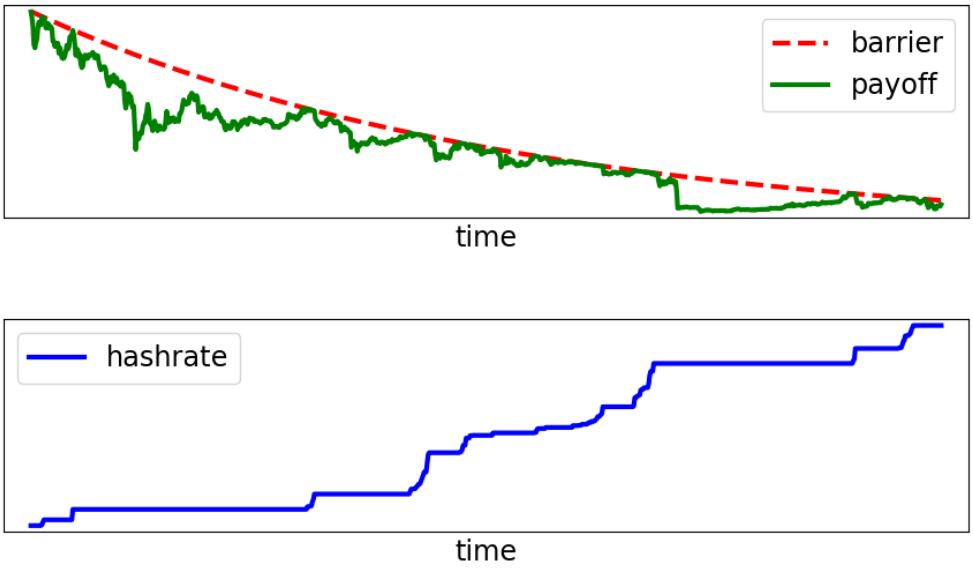
Proposition 1.2. *Assume that assumptions 1, 2, 3 and 4 hold. Then there is an industry equilibrium* $\left(P_t, \overline{P}_t\right)$ *such that* P_t *is a GBM reflected at* $\overline{P}_t = \overline{P}_0 / A_t$ *where*[14]

$$\overline{P}_0 = \frac{\beta(r-\alpha)}{\beta - 1}\left[I_0 + \frac{C_0}{r}\right], \text{ and } \beta = \frac{\frac{\sigma^2}{2} - \alpha - a + \sqrt{\left(\alpha + a - \frac{\sigma^2}{2}\right)^2 + 2\sigma^2(a+r)}}{\sigma^2} > 0. \tag{7}$$

A typical equilibrium is illustrated in Figure 1. The upper-panel reports an arbitrary sample path for the payoff process $(P_t)_{t\geq 0}$. Payoffs follow the changes in the exchange rate and thus behave as a GBM until they hit the reflecting barrier $\overline{P}_t$. Such events trigger market entry, as shown in the lower-panel. The resulting increase in hashpower raises the difficulty of the mining problem and thus pushes payoffs down until market entry is not anymore profitable. The entry barrier decreases at the rate of technological progress because it corresponds to the pace at which both entry and operating costs fall over time.

[14] Note that, when $\alpha = r$, $\overline{P}_0 = \left(I_0 + \frac{C_0}{r}\right)\left(\alpha + a + \frac{\sigma^2}{2}\right)$.

Figure 1: Industry Equilibrium



Comparative statics.—The higher the barrier, the lower the average rate of investment as miners procrastinate further before entering the market. It is therefore instructive to study the effect of the parameters on $\overline{P}_0$. Differentiating its expression in (7), we find that $\partial\overline{P}_0/\partial a > 0$ and $\partial\overline{P}_0/\partial r > 0$. If technological progress accelerates, miners' revenues shrink more rapidly because there will be more entries in the future. Hence miners have to earn more in the periods following their entries and so the barrier must be higher. A similar mechanism explains the impact of r since the value of future profits is discounted at a higher rate when r goes up. Not surprisingly, an increase in the average growth rate α of the block reward incentivizes entry as $\partial\overline{P}_0/\partial\alpha < 0$. Finally, the volatility of payoffs σ discourages entry since $\partial\overline{P}_0/\partial\sigma > 0$. Note that this effect is not due to an increase in the value of waiting because the perfectly competitive structure of the industry rules out such an option: competitors would preempt any procrastination beyond the zero expected profit threshold. Instead, the negative impact of σ on $\overline{P}_0$ is mechanical. Given that payoffs are truncated from above by the reflecting barrier, an increase in their spread automatically lowers their expected value. Quantitatively, the rate of technological progress a has, by far, the largest effect on $\overline{P}_0$.

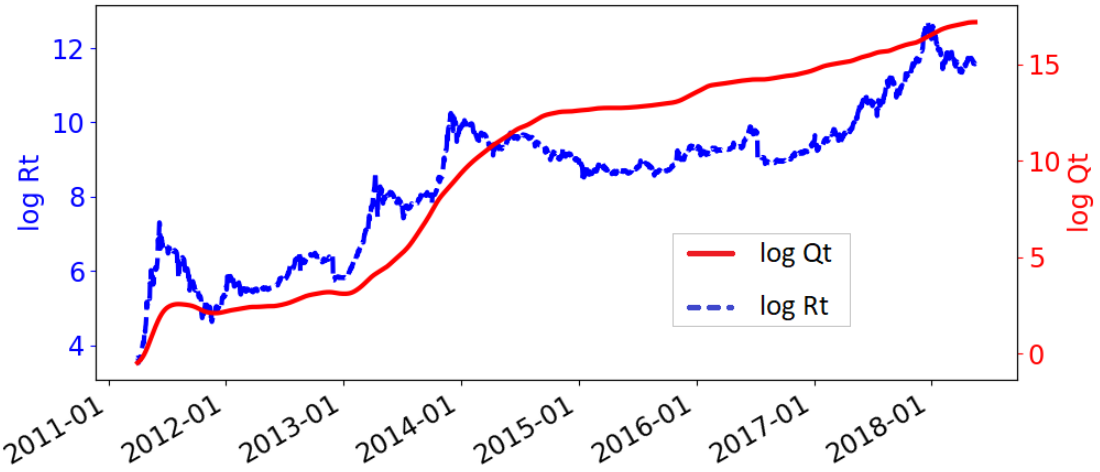
1.4 Calibration

Data.—We now show that feeding our model with exchange rate data allows one to accurately predict the evolution of the network hashrate. For this purpose, we need to infer the miners' payoffs $P_t = R_t/Q_t$. Remember that the numerator, R_t, is equal to the value of new coins plus the transaction fees. The number of created coins per block is specified by the protocol while the exchange rate is directly available from coindesk.com.[15] The transaction fees are recorded in the blockchain and can easily be retrieved from btc.com. Thus all the components of $(R_t)_{t\geq 0}$ are readily available. This is, however, not the case for the network hashrate $(Q_t)_{t\geq 0}$ whose values must be estimated using the theoretical probability of success and the number of blocks found each day. Given that we are not primarily interested in statistical inference, we relegate the description of our estimation procedure to Appendix 1.7 and save on notation by using Q_t to denote our estimate, although its time series only approximates the true hashrate. We show in Appendix 1.7 that the approximation is accurate. We update the value of Q_t on a daily basis and, since there are on average 144 blocks mined every day, the expected payoffs per day are given by $P_t = 144 \times R_t/Q_t$.

We report the series followed by $(R_t)_{t\geq 0}$ and $(Q_t)_{t\geq 0}$ in Figure 2. There is a clear correlation between the two variables. Our model suggests that their structural relation should become apparent if one takes the ratio of the two series and detrend it at the rate of technological progress a. Then the resulting series should behave as a reflected Brownian motion. A natural guess for the rate of progress is Moore's law according to which processor speed doubles every two years. We actually expect improvements in the mining technology to outpace those in processing speed because miners came up with a series of innovations which allowed them to leverage their computing power. Thus we will refine our guess later on by calibrating the value of a. Yet it is instructive at this exploratory stage to use Moore's law as a benchmark.

[15] There are many different exchanges and the exchange rates vary a bit across them. We neglect those variations since they are dwarfed by the changes over time of the exchange rate.

Figure 2: Miners Revenues R and Network Hashrate Q

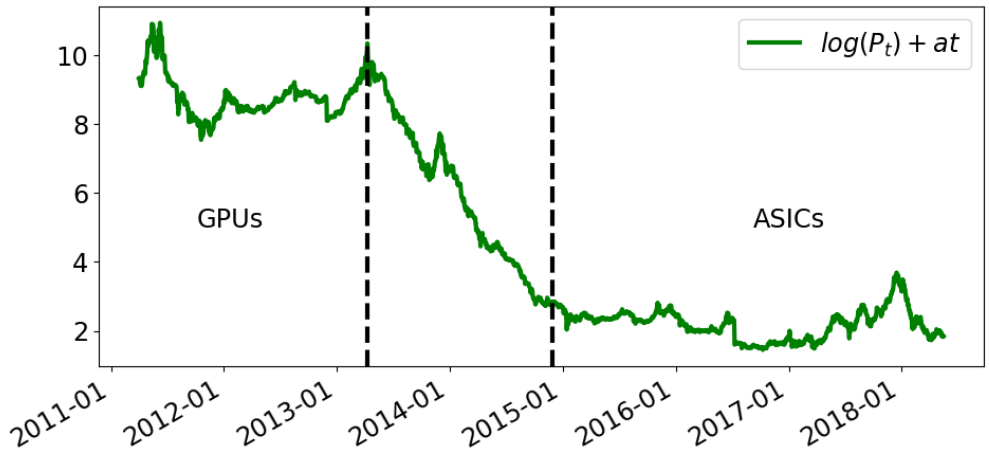


Note: R_t is computed using information collected on coindesk.com and btc.com. Q_t is inferred using the procedure described in Appendix 1.7.

The detrended payoff series based on Moore's law is reported in Figure 3. It exhibits two stationary regimes, with a break in the middle where payoffs decreased regularly until they reached a lower plateau. At first, this behavior does not seem to square with the model. But if we focus on the date at which the break initiates, we realize that it coincides with the switch to Application Specific Integrated Circuits (ASICs). Since this revolution in the mining technology boosted the rate of progress well above its long-run trend, Assumption 4 does not hold and thus one should not expect the predictions of our model to be verified during this transitory phase. Hence we leave aside the lapse of time where miners switched from GPUs and FPGAs to ASICs, and focus instead on the two subperiods where miners used the same technology. More precisely, during the first period, which ranges from 2011/04/01 to 2013/01/31, miners mainly mined with GPUs; while they mostly relied on ASICs from 2014/10/01 onwards. Our second subperiod ranges from 2014/10/01 to 2017/03/31. So far, we leave aside the most recent period, which witnessed the birth and death of a giant bubble in the ₿/$ exchange rate, because the model is unable to account for those data. We will, however, make clear below why this very period is problematic and why it should not be interpreted as evidence against the accuracy of our model. Note that the first halving of the monetary creation rate happened on 2012/11/28, towards the end of the first period, while the second halving happened in 2016/07/09, around the middle of the

second period.

Figure 3: Detrended Payoff Series



Note: P_t has been computed dividing $144R_t$ (the daily network revenue) by Q_t.

Calibrating the parameters.—We calibrate the parameters for each subperiod. The model is parsimonious enough to rely on six parameters only: the deterministic trend α of rewards and their volatility σ^2, the rate of technological progress a, the discount rate r, the price I_0 of one unit of hashpower bought at time 0 and the electricity cost C_0 of that same unit. The first two parameters can be directly estimated using $(R_t)_{t\geq 0}$ only. Under assumption 3, the log returns are independent and follow a normal distribution with mean $\mu \equiv \alpha - \sigma^2/2$ and variance σ^2, which we estimate by maximum likelihood (see Appendix 1.7).

The rate of technological progress, a, and the reflecting barrier, $\overline{P}_0$, are set to minimize a (pseudo)distance between the observed and the simulated paths of the hashrate. A direct consequence of our equilibrium definition is that $Q_t = \max\left(Q_{t-1}, \frac{R_t A_t}{\overline{P}_0}\right)$ for all t. This condition provides us with a straightforward way to simulate the hashrate for any sample with T observations:

1. Set the initial value of the simulated hashrate Q_0^{sim} equal to its empirical counterpart, i.e. $Q_0^{sim} := Q_0$.
2. Update the simulated hashrate as follows: $Q_t^{sim} := \max\left(Q_{t-1}^{sim}, \frac{R_t A_t}{\overline{P}_0}\right)$, for $t = 1, \ldots, T$.

Since $(R_t)_{t\geq 0}$ and Q_0 are observed, the minimization procedure boils down to finding the value of a and $\overline{P}_0$ such that

$$(\hat{a}, \hat{\overline{P}}_0) \in \underset{(a,\overline{P}_0)\in\mathbb{R}^+\times\mathbb{R}^+}{argmin} \sum_{t=1}^{T} \left(\frac{Q_t - Q_t^{sim}(a, \overline{P}_0)}{Q_t} \right)^2. \tag{8}$$

Unfortunately, the three other parameters $\{r, I_0, C_0\}$ cannot be disentangled. We therefore fix r, and recover the total costs of one terahash per second bought at the beginning of each subperiod, $K_0 \equiv I_0 + C_0/r$, by equating the expression for $\overline{P}_0$ in (7) with the estimated $\hat{\overline{P}}_0$. The arbitrary choice for the discount rate r turns out to be relatively unimportant because the term $\beta(r-\alpha)/(\beta-1)$ in (7), and thus total costs, are rather inelastic with respect to r.[16] The parameters resulting from our calibration strategy are summarized in Table 1, where all values are expressed as yearly rates.[17]

Table 1: Calibrated Parameters

Method	Parameter	Interpretation	1st period	2nd period
(fixed)	r	Discount Rate	0.1	0.1
(estimated)	μ	Trend $log(R_t)$	1.41	0.19
	σ^2	Variance $log(R_t)$	1.95	0.54
(calibrated)	a	Rate of TP	1.18	0.76
	K_0	Total Costs	$\$5.6 \times 10^6$	\$ 1825

According to Moore's law, a should be close to $\log(2)/2 \approx 0.35$ since it predicts that the price of one unit of hashpower is divided by two every two years. Our calibration suggests that the mining technology progressed at a much faster rate although it slowed down considerably in the second period. This finding is consistent with the conjecture that miners were able to implement innovations specific to the hashing problem on top of the raw increase in computing power. But such improvements became harder to unearth as the mining technology matured and the rate of progress gradually converged towards the one predicted by Moore's law.

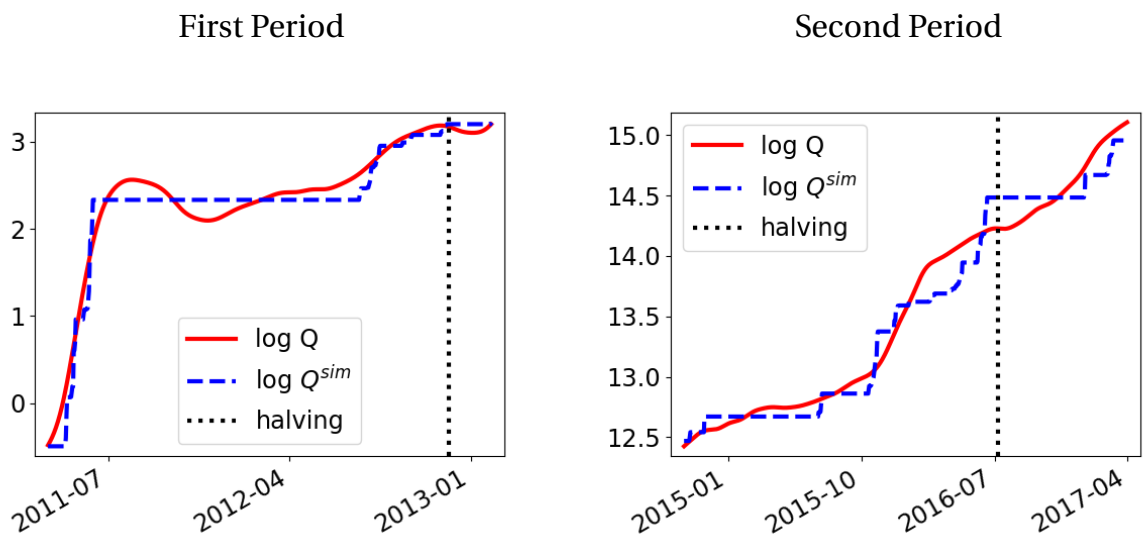
The average growth rate of rewards, μ, also decreased a lot between the two periods of study. As one would expect, early buyers of Bitcoins earned higher returns.

[16] In the second period, setting $r = 0.2$ yields $K_0 = \$\,1639$, while $r = 0.05$ yields $K_0 = \$\,1934$.

[17] For example, the estimates for a means that the price of a new machine has been on average divided by $\exp(a)$ every year during each period of study.

Information about their profits pushed the demand for Bitcoins which raised the exchange rate even more. But the extremely high returns observed at the beginning became harder to sustain as the market capitalization grew from a negligible amount to nearly 20 billions $ by the end of our sample. In spite of this cooling process, investing in Bitcoin remained extremely profitable, especially when one bears in mind that the values we report for μ take the halvings into account. These tremendous returns have led many observers to announce the imminent collapse of Bitcoin.[18] Whether or not such predictions will eventually be vindicated is beyond the scope of this paper, but our estimates for the volatility coefficient σ indicate that there was no obvious arbitrage opportunity as investors willing to bet on Bitcoin also had to bear a huge risk. Even though the volatility of rewards was divided by three in the second period, its value remained an order of magnitude higher than its counterpart for the S&P 500.[19]

Figure 4: Simulated vs Observed Hashrates



Predicted vs actual hashpower.—The estimation procedure provides us with an estimate for the reflecting barrier, $\overline{P}_0$, as well as for its trend, a. Using these two values, we can run the two-step algorithm described above to simulate the network hashpower $(Q_t^{sim})_{t\geq 0}$. We report the simulated series against its empirical counterpart in Figure 4. In spite of its very parsimonious structure, the model tracks the actual hashpower remarkably well over the long run. We nonetheless notice some temporary dis-

[18] According to bitcoinobituaries, by May 2018, 299 opinion pieces had already predicted the death of Bitcoin.

[19] We find that, for the S&P 500, $\sigma^2 = 0.053$ for the first period and $\sigma^2 = 0.027$ for the second period

crepancies. In particular, during the second period, the model is a bit less accurate around the halving date (2016/07/09). This is not surprising because miners do not anticipate halvings in our model while they certainly do in reality. Hence, it is actually more intriguing that such a disconnect between the simulation and the data is not apparent around the first halving date (2012/11/28). The explanation is the following: the technical progress rate was so high during the first period that miners' payoff were anyway very low at the time of the halving, except for those who entered the race just before. We investigate this conjecture in Section 1.5.1 where we explicitly introduce halvings into our model.

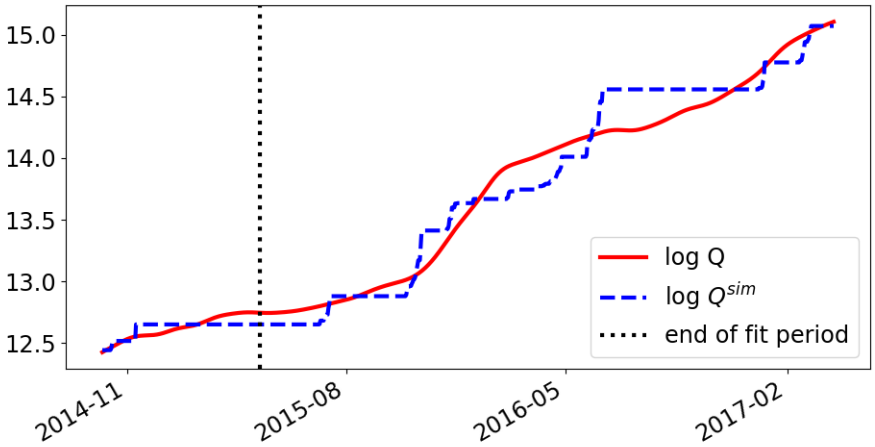
Another noticeable difference between the actual and the simulated hashrates is that the former sometimes decreases, especially during the first period, while the latter never does. Our model cannot reproduce these drops in haspower because it is based on the premise that investment is totally irreversible. We will address this shortcoming in Section 1.5.2 by allowing miners to mothball or scrap their machines.

These discrepancies do not invalidate our approach since the model was devised to capture long run trends in haspower. Yet one could argue that such a conclusion is too generous because our procedure minimizes the distance between the simulation and the data, and so would fit the data fairly well even if the model were misspecified. Our first answer to this argument is that we optimize on two parameters only, which is not much to fit times series of 608 and 913 data points. Moreover, to perform the simulations we start from the initial hashrate for each subperiod and then let the model run without using intermediate realizations to correct its output. Given that our estimation procedure does not place any additional weight on the final values of the hashpower, any fundamental misspecification would have generated a noticeable gap between the simulation and the data during some subperiod. Thus we view the fact that there is no obvious deterioration of the model's fit over time as a convincing verification of its accuracy. We now provide support for this interpretation by performing out-of-sample tests, and by comparing the entry rule predicted by the model with the one prevailing in the data.

Out-of-sample tests.—We assess the model's ability to match out-of-sample data by dividing the second period into a fit period and a test period. We calibrate a and $\overline{P}_0$ on the fit period only and find that, even when the fit period is pretty small, the calibrated values remain close to the ones based on the full sample. Thus the predicted hashrate remains accurate several years after the end of the fit period, as shown in Figure 5.

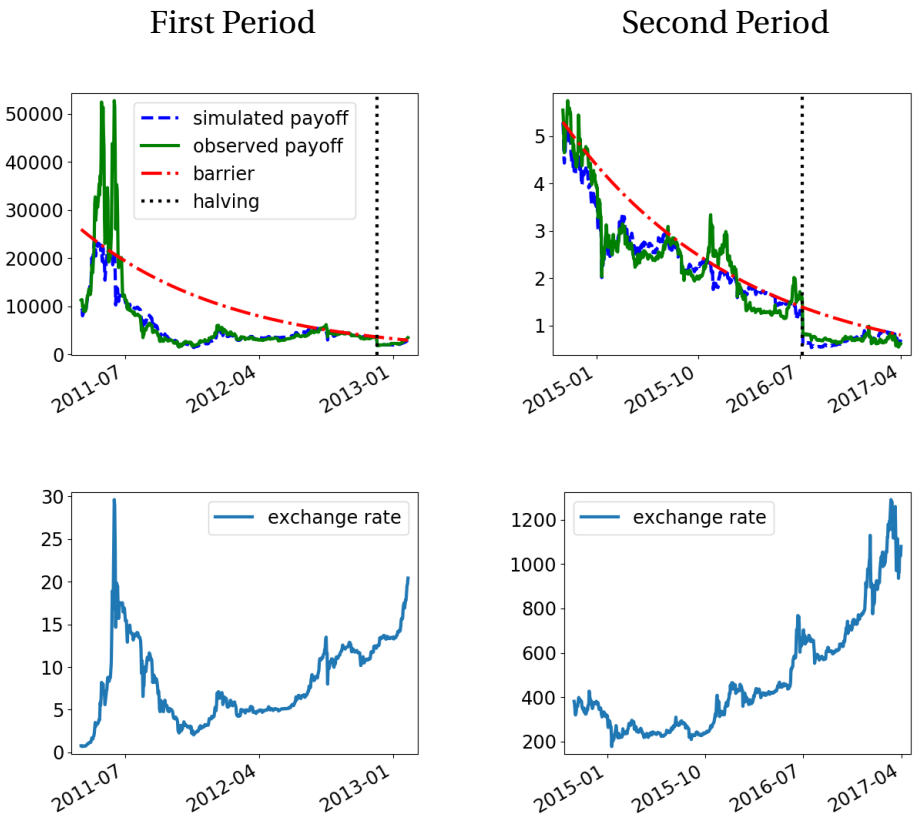
Note, however, that out-of-sample tests are much less conclusive for the first sub-period because the hashrate increases only at the beginning and at the end of that period. Hence, if we split the first data sample into a fit and a test period, the payoffs do not hit the reflecting barrier often enough to deliver reliable calibrations. For instance, the parameters are not identified if the payoffs hit the barrier only once as one cannot pinpoint a line with the help of a single point.

Figure 5: Out-of-sample Test



Note: The fit period is the shortest one for which the overall fit remains accurate.

Figure 6: Simulated vs. Observed Payoffs



Inspecting the entry rule.—Besides assessing the model's overall fit, we can also check whether the data are in line with the s-S rule predicted by the theory. For this purpose, we report the simulated and observed payoffs in the upper-panels of Figure 6. As forecasted by the model, the observed payoffs remain below the barrier most of the time and tend to reflect downwards when they reach its vicinity. This is remarkable in itself since $\overline{P}_t$ was estimated regardless of this requirement, fitting the hashrate only.

Although the observed and simulated hashrate series are nearly superimposed for most of the dates, there are some time intervals where the two series differ significantly. These divergences occur for two reasons. First, the fit of the model deteriorates significantly around the halving date (2016/07/09) of the second period. But, as explained above, this is precisely what one should expect since the model does not take halvings into account. Second, the model sometimes fails due to extreme realizations of the exchange rate. This can be seen by comparing the upper-panels of Figure 6 with the lower panels where we report the ₿/$ exchange rate. One quickly notices that the

periods of divergence between observed and simulated payoffs are clustered around the dates where the exchange rate is extremely volatile. Quite intuitively, when the exchange rate goes up 30% or more in one day,[20] miners cannot enter the market as quickly as the model predicts because they are facing, among many other frictions, delivery and manufacturing delays. Devising a model that takes into account such constraints, by introducing frictions along with potentially convex adjustment costs at the industry level, would probably improve the correspondence between the two series. We leave such refinements to further research because they greatly complicate the solution of the model,[21] while our findings suggest that they are not likely to yield significant forecasting gains beyond short-term horizons.[22]

The 2017 bubble.—While this paper was being written, Bitcoin experienced a period of trading frenzy. From $\$3,226$ on the 14th of September of 2017, Bitcoin's exchange rate shot up to $\$19,343$ on the 16th of December and then dropped back to $\$6,914$ on the 5th of February of 2018.[23] Since then, the exchange rate has somewhat recovered and fluctuated around $\$9,000$. Perhaps not surprisingly, our model indicates that the relation between the exchange rate and the network hashrate broke down during that period. Figure 7 shows that, if the relation had remained stable, the hashrate should have been six times higher at the peak of the bubble.[24]

The discrepancy between the predicted and observed hashrate is explained by three different factors. First, investment in hashpower was constrained by delivery delays. In May 2017, there were approximately 230,000 active machines. Between May and December 2017, ₿/$ exchange rate was multiplied by 12. To keep up with this pace, approximately 2,700,000 new machines would have had to be installed within eight months only. Such a dramatic increase was bound to stretch the productive capacity of Bitmain, the manufacturer of ASICs for Bitcoin mining. Second, Bitmain being the only producer of ASICs, it probably exercised his monopoly power and decided not to flood the market with new machines in order to raise their selling price. Indeed,

[20] Extreme daily gains of 30% or more were observed on 05/10/2011, 06/03/2011, 04/17/2013, 11/18/2013 and 12/18/2013.

[21] See for example the work of Aid et al. (2015) on regulated Brownian motions with delays.

[22] This conjecture is supported by the observation that, when the payoff variable temporarily exceeds the barrier due to a surge in the exchange rate, it tends to quickly decrease afterwards. These corrections are very much in line with our model: they occur because the hashrate catches up and not because the exchange rate decreases.

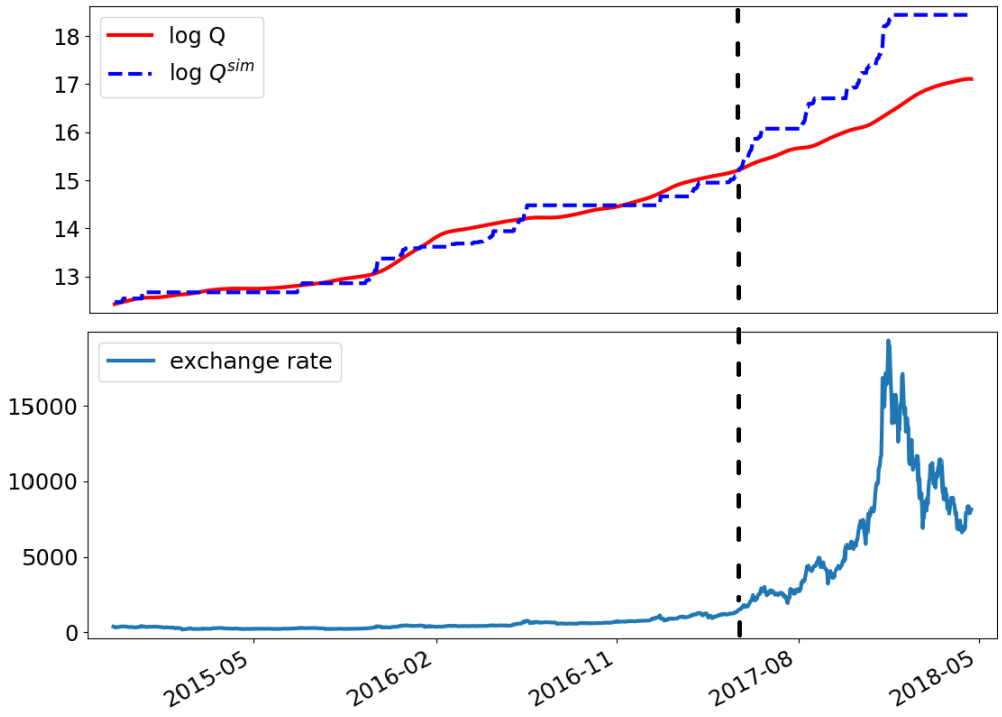
[23] Exchange rates as reported by coindesk.

[24] Remember that Figure 7 uses a log-scale for the hashpower.

the price of an Antminer S9 mining rig was multiplied by three between the beginning and the climax of the bubble, and then divided by around four during the following crash.[25] Third, it seems plausible that some miners waited to see whether or not the boom was sustainable, and so wisely refrained from over-investing in mining power.

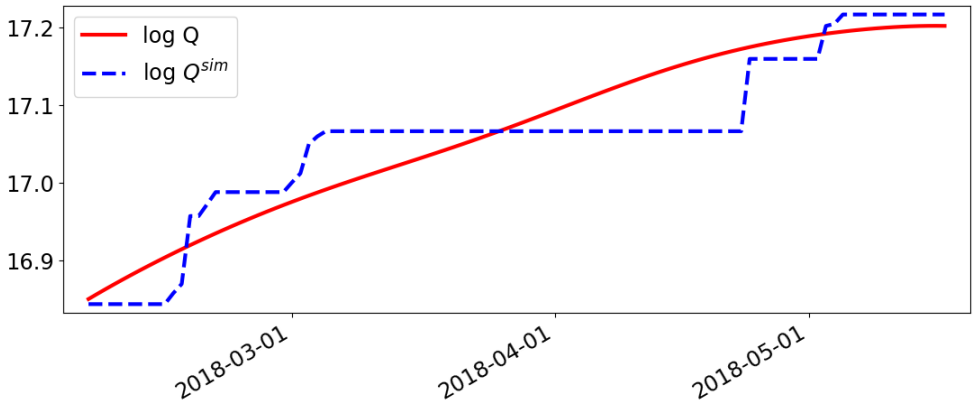
A supporting evidence for our interpretation is that, as of April 2018, the hashrate would have caught up with its predicted value if the bubble never happened. Looking at the lower panel of Figure 7, one sees that the exchange rate in November 2017 is the same as in April 2018; while the upper panel shows that the value of the predicted hashrate in November 2017 is also equal to the value of the observed hashrate in April 2018. This is consistent with delivery delays of five to six months which would have prevented miners from investing at the peak of the bubble. Of course, our model does not take into account such constraints and so overestimates the hashpower after the bubble burst. However, our model is again accurate if we reinitialize its calibration at the end of the bubble (see Figure 8), thus providing further evidence that its failure during the bubble is probably explained by manufacturing delays and congestion externalities.

Figure 7: The 2017 bubble



[25] See https://www.anythingcrypto.com/guides/bitcoin-antminer-S9-prices-2018.

Figure 8: After the bubble



1.5 Extensions

Our model hinges on two simplifying assumptions: (i) the rate of money creation is kept constant; and (ii) miners do not have the option to turn off their machines. We now relax each one of them and describe the extent to which such generalizations improve the fit of the model.

1.5.1 Model with halvings.

So far we have assumed that revenues follow a GBM (Assumption 3). Thus we have ignored that the number of new coins issued per block is divided by two every 210,000 blocks. These so-called halvings create discontinuities in the paths of R_t that are inconsistent with the GBM specification. In this subsection, we take halvings into account by replacing Assumption 3 with Assumption 5, so that block rewards are divided by two every four years.

Assumption 5. *The block reward is equal to* $R_t = h_t \tilde{R}_t$. $\tilde{R}_t$ *follows a GBM while* $h_t = \left(\frac{1}{2}\right)^{\lfloor t/4 \rfloor}$, *where* t *measures the number of years elapsed since the inception of Bitcoin.*

Assumption 5 slightly simplifies the halving process. First, the reward a miner gets when she finds a block is not exactly divided by two after each halving because it includes transaction fees on top of new coins. The discrepancy is, however, not very important in our data sample since we study only the first two halvings and, by the time of the second halving, transaction fees always accounted for less than 2% of av-

erage block rewards.[26] Second, halvings do not occur every four years, but instead every 210,000 blocks. Counting years is a way to approximate elapsed time because the Bitcoin protocol adjusts the difficulty of the hashing problem every 10 minutes on average.[27] Appendix 1.7 shows that the updating rule managed to keep the block-finding rate close to one every 10 minutes.

Halvings render the miners' problem non-stationary: the closer they are to the halving date, the lower their expected payoffs. Hence we cannot anymore solve for the entry barrier in closed-form. Instead, we have to rely on numerical methods. We proceed by backward induction: starting from the stationary solution derived in the previous section, we use a finite-difference procedure to approximate the entry rule. Going back in time, the algorithm quickly converges towards an entry barrier that is independent of the number of future halvings.[28]

The values of the non-calibrated parameters $\{r, \mu, \sigma\}$ are the same as in the baseline model. The numerical procedure described above allows us to recover the entry barrier for given values of a and K_0, thus enabling us to use the algorithm outlined in Section 1.4 to simulate the network hashrate. Minimizing the distance between the simulated and the observed paths yields the parameter values reported in Table 2.

The introduction of halvings has a negative impact on mining costs K_0 and a positive one on the rate of technological progress a. The increase in K_0 is quite intuitive: since halvings decrease expected revenues, free entry requires that mining costs decrease too. The reason why a increases is a bit more subtle. This adjustment corrects the misspecification of the baseline model which necessarily overestimates the hashrate around the halving dates. This is why the minimization procedure, when applied to the baseline specification without halvings, generates a negative bias for a because it uses its value to reduce the discrepancies around the halving date.

[26] There are a few exceptions when some users mistakenly sent huge amount of transaction fees. For instance, on 2016/04/26, a transaction gave 291 bitcoins as fees.

[27] Note that, in our model, the block-finding rate is constant since we assume that the difficulty of the mining problem is updated continuously (see Assumption 1). Hence Assumption 5 could equivalently be stated using the numbers of created blocks instead of calendar time.

[28] More precisely, the entry barrier is essentially stable after four iterations. We use finite-difference methods to approximate the Hamilton-Jacobi-Bellman equations satisfied by the value functions of miners. We rely on the implicit Euler scheme in order to ensure that the approximation is stable. The system of linearized equations is solved using a generalization of the Gauss-Seidel iterative method known as the successive-over-relaxation method.

Table 2: Calibrations with and without Halvings

Parameter	Interpretation	1st period		2nd period	
		Halvings	No Halvings	Halvings	No Halvings
a	Rate of TP	1.29	1.18	0.85	0.76
K_0	Total costs	$\$5.3 \times 10^6$	$\$5.6 \times 10^6$	\$ 1,655	\$ 1,825

Note: all the other parameters are as reported in Table 1.

Note, however, that these corrections are rather small, which should not be surprising since the baseline model was already pretty accurate. Figure 9 shows that, for the first period, the paths predicted by the models with and without halvings are nearly identical. As conjectured in the previous section, neglecting the first halving was not so important due to the extreme volatility of the exchange rate at that time. For the second period, the simulated paths remain very close to each other except around the halving date where, as one should expect, the extended model outperforms the baseline specification. This indicates that the halving affected miners' behavior only a couple of months ahead, a conjecture which can be substantiated by looking at the entry barrier.

Figure 9: Hashrates Comparison

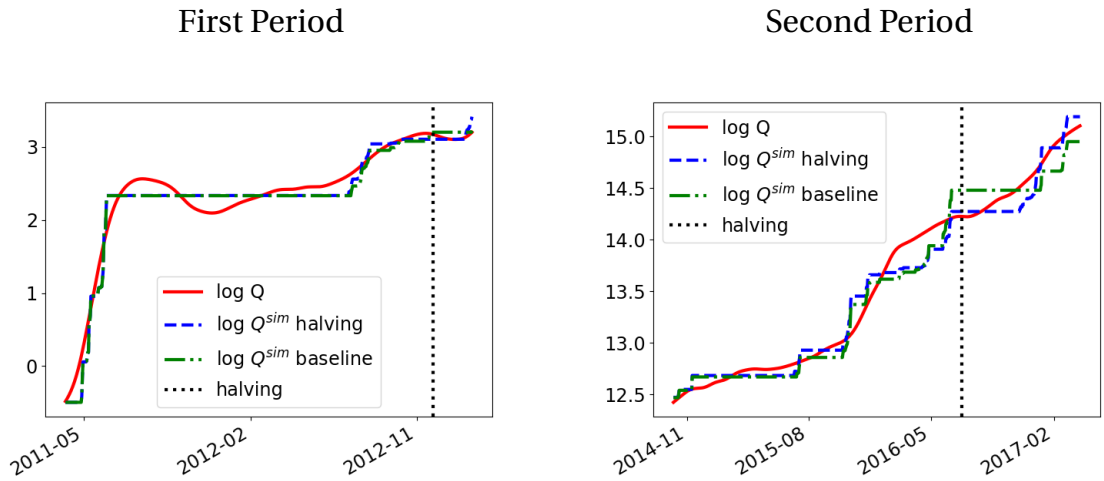
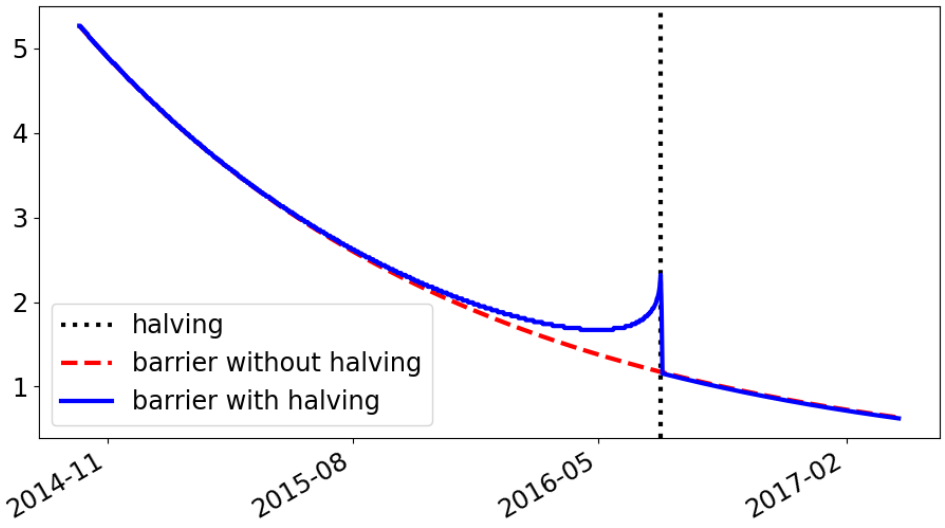


Figure 10 displays the shape of the entry barrier in the second period.[29] By construction, the barrier shifts down by 50% on the day of the halving. This drop follows

[29]The barrier obtained for the first period is very similar and thus not displayed.

a period where the barrier slopes up because miners anticipate the drop in future revenues and so procrastinate more before entering the market. But this adjustment becomes noticeable only a few months before the halving and is therefore not relevant for most of the period. This might be surprising given that a division by two of revenues seems like a huge loss. Two effects can account for this fact. First, miners who enter the race a couple months before the halving enjoy anyway, when this event occurs, a payoff substantially lower than at their entrance times. Second, for the same miners, comparatively to the model without halving, the future loss of income is partially compensated by a higher revenue before the halving, due to a lower hashrate. When the technical progress rate decreases, halvings will have a more noticeable impact on the network hashrate.

Figure 10: Entry Barrier with Halvings



1.5.2 Mothballing and scrapping options.

We now relax Assumption 2 according to which miners always keep their hardware in mining mode. In practice, miners have the option to switch off their machines, and they can switch them back on should mining become profitable again. We assume that the hardware can be kept idle at zero costs. Thus the mothballing decision is fully reversible, and as such does not involve any forward-looking component. Machines are mining whenever their flow revenues are higher than their operating costs.

In other words, the per-period profits at time t of a miner entered at time τ are equal to $\max(P_t - C_\tau, 0)$, and their value functions read

$$V(P_t, \tau) = \mathbb{E}_t \left[\int_t^{+\infty} \max(P_s - C_\tau, 0)\, e^{-r(s-t)} ds \right]. \tag{9}$$

If there is no technological progress, all miners pay the same electricity costs (C_τ is constant) and thus face the same problem. Then the industry equilibrium features two reflecting barriers: an upper-barrier generated by the entry of new miners which push payoffs downwards until free entry is satisfied again, and a lower barrier generated by the exit of incumbents which push payoffs upwards until miners are indifferent between operating and stopping their machines.[30] With technological progress, the structure of the industry is much more intricate. Then miners cannot all be indifferent since they bear different operating costs. The least productive miners are the first to mothball their machines, and they do so until the marginal miner makes zero flow profits. This endogenous cutoff depends on the distribution of vintages among incumbents. Thus the law of motion of P is not anymore a function of current revenues only, but also of the vintage distribution. This in turn greatly complicates the decision of prospective entrants who now have to solve a problem which includes a distribution function among its state variables.

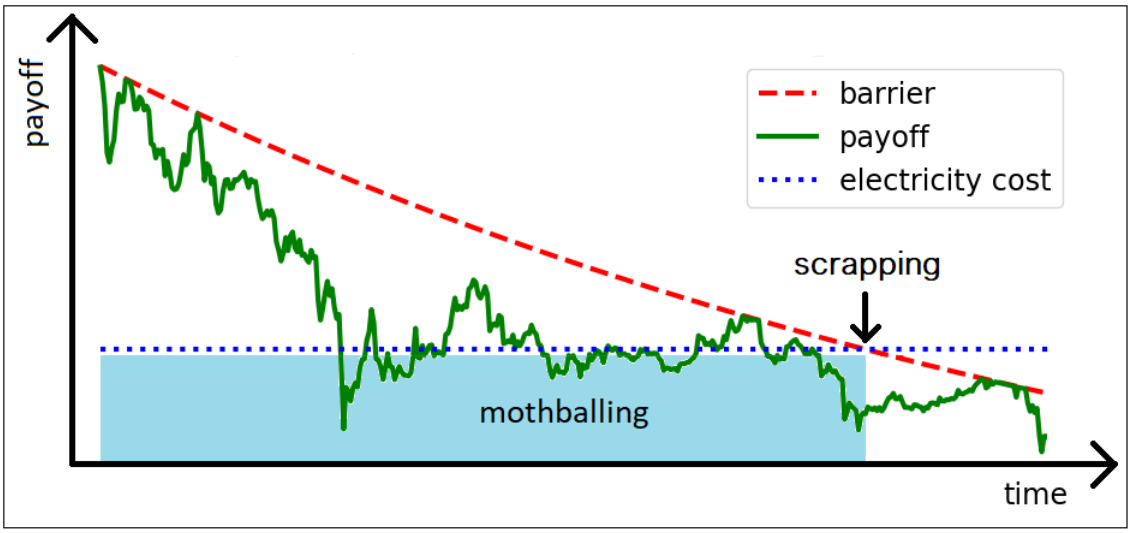
Analyzing such a problem involves devising new and complex numerical methods. Instead of following this direct approach, we take the view that prospective entrants do not have access to the data required to solve the full information problem. Finding the vintage of all machines is an extremely tedious, if not impossible, task. It is therefore quite unlikely that miners actually looked for this information before investing and, even if they did, they would only have observed a very noisy measure of the actual distribution. We assume instead that potential entrants make their decisions considering only the current value of the flow payoffs. We establish below the plausibility of this restriction by showing that mothballing and scrapping have very little impact on the hashrate, so that entrants cannot significantly benefit from solving the full information problem. From a formal standpoint, we assume that miners' expectations satisfy the following Markov property.

Assumption 6. *Let $\mathcal{F}_t \equiv \sigma(P_s; 0 \leq s \leq t)$ denote the filtration generated by P. We assume that, for all measurable set $A \in \mathbb{R}_+$ and all $s > t$, $\Pr^e(P_s \in A|\mathcal{F}_t) = \Pr^e(P_s \in A|P_t)$,*

[30] See Alvarez and Shimer (2011) for a model with two reflecting barriers generated by workers entry and exit.

where $\Pr^e(\omega)$ *is the probability of event* ω *as evaluated by potential entrants.*

Figure 11: Mothballing and scrapping regions



We show in the proof of Proposition 1.3 that, under Assumption 6, the equilibrium is again characterized by an entry barrier $\overline{P}_t$ which decays at the rate of technological progress. Since payoffs are reflected downwards when they hit the barrier, it will never be profitable to operate a piece of hardware which is so obsolete that its operating costs exceed the entry barrier. A typical mining cycle is illustrated in Figure 11: the machine is mothballed whenever payoffs fall below its operating costs, as indicated by the colored area; and it is scrapped when the entry barrier crosses the operating costs. The addition of this exit threshold makes it impossible to analytically solve for $\overline{P}_0$.

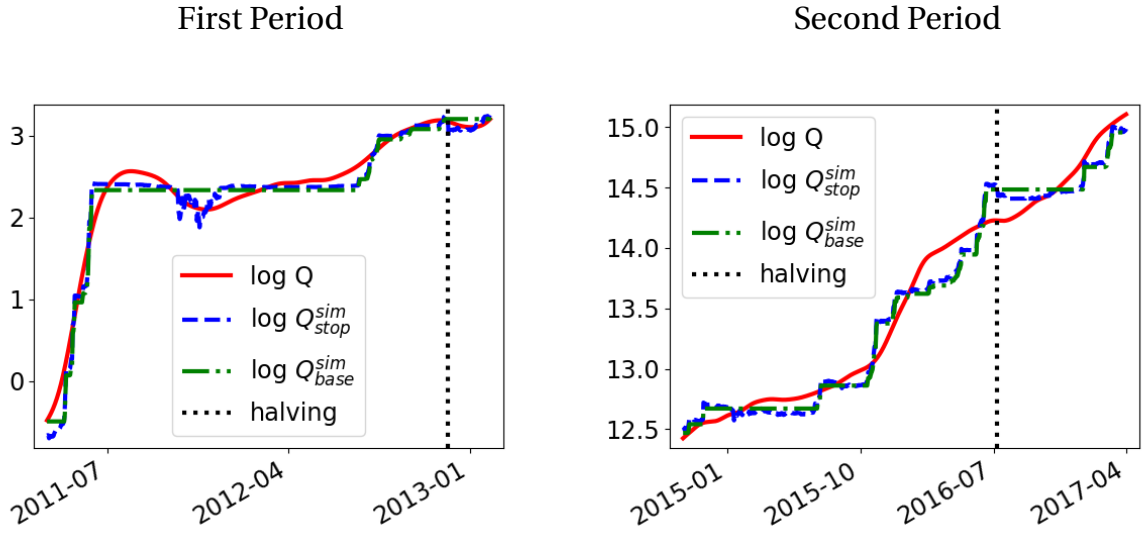
Proposition 1.3. *Assume that Assumptions 1, 3, 4 and 6 hold true. Then there exists a* $\overline{P}_0 > 0$ *such that* $\left(P_t, \overline{P}_t = \overline{P}_0/A_t\right)$ *is an industry equilibrium that satisfies the requirements of Definition 1.1.*

Simulating the hashrate.—Simulating the hashrate is more complicated than for the baseline model because one must keep track of the electricity costs, as well as of the activity status of all miners. The inputs of the algorithm are the exchange rate $((R_t)_{0\leq t\leq T})$, the rate of technological progress a, and the initial hashrate, electricity costs and entry barrier $\left(Q_0, C_0, \overline{P}_0\right)$. We also need to initialize the vintage distribution of all active miners. A series of robustness checks demonstrates that the initial choice

of vintages hardly affects the simulated paths after a couple of days.[31] This is in line with Assumption 6 since knowing the true distribution of miners' vintages does not significantly improve forecast accuracy.

The simulation procedure works as follows. For each day, we start by deleting from the database all miners whose electricity costs are bigger than the entry barrier. As explained before, those miners scrap their machines because they will never find it profitable to mine in the future. Then, given the new value of R_t, we compute the temporary payoff miners would face if the hashrate remained constant. Depending on its value, two configurations may arise. First, if this temporary payoff is smaller than the electricity costs of the least efficient (i.e. oldest) active miner, we know that some miners should switch off their machines. Thus we let the least efficient miners mothball their hardware, and update the temporary payoff until no active miner prefers to remain idle. Alternatively, if the temporary payoff is higher than the electricity costs of the most efficient inactive miner, we know that some miners should switch on their machines. Finally, if all incumbents are active and the temporary payoff is still bigger than the entry barrier, we let new miners enter the market until the temporary payoff equals the entry barrier.

Figure 12: Simulated vs Observed Hashrates



[31] We therefore pick a distribution of vintages for which the mass of miners of any vintage is inversely proportional to the price of their vintage, as would have happened if the environment were deterministic.

Predicted vs actual hashpower.—As before, we calibrate a, $\overline{P}_0$ and C_0 so as to minimize the distance between the simulated and actual hashrates. Figure 12 reports the resulting series along with the prediction of the baseline model. As expected, the model with exit fits the data better whenever the actual hashrate decreases. But the improvement is rather marginal because significant decreases in hashrate are exceptional events. Most of the time the predictions of the two models coincide, thus substantiating our claim that Assumption 2 is a reasonable benchmark.

Disentangling investment from operating costs.—We now add an assumption which allows us to disentangle the price of machines from their operating costs. The simulations reported in Figure 12 show that entrants can ignore the impact that mothballing and scrapping have on the hashrate, and nonetheless make accurate predictions about their future payoffs. Hence we strengthen Assumption 6 and let entrants disregard the rare instances where the hashrate shrinks.

Assumption 7. *When forming their expectations, potential entrants ignore the impact that mothballing and scrapping have on the network hashrate.*

Entrants who base their expectations on the premise that incumbents will never stop mining can disregard the technology operated by other miners. Thus Assumption 7 implies Assumption 6, although the converse is not true. Assumption 7 ensures that expected payoffs follow a reflected GBM as in the baseline model. Knowing the distribution of P allows us to compute the entrants' expected profits. In particular, equation (9) is compatible with free entry if and only if

$$I_t \geq \int_t^{+\infty} \left(\int_t^{\overline{P}_s} \max\left(y - C_t, 0\right) f^e_{P_s|P_t=\overline{P}_t}\left(y\right) dy \right) e^{-r(s-t)} ds, \text{ for all t,} \tag{10}$$

where $f^e_{P_s|P_t=\overline{P}_t}(\cdot)$ denotes the distribution of P_s conditional on $P_t = \overline{P}_t$ as anticipated by entrants. For the sake of conciseness, we defer the explicit expression of f^e to Lemma 1.1 in the Appendix. Evaluating the integral on the right-hand side for the calibrated values of a, $\overline{P}_0$ and C_0 yields the consistent investment cost I_0. Equation (10) also places an upper bound on total costs paid by miners. Let T denote the time it takes for the entry barrier to reach the electricity costs of new entrants. Given that the entry barrier decays at the rate of technological progress, we have $T = \log\left(\overline{P}_0/C_0\right)/a$. The total costs paid by miners who entered at time 0 must necessarily be inferior to $K_0 = I_0 + \int_0^T C_0 e^{-rt} dt$ because they will never find it profitable to mine after date T.

Table 3: Calibration with and without Exit

Parameter	Interpretation	1st period		2nd period	
		Exit	Baseline	Exit	Baseline
a	Rate of TP	1.15	1.18	0.76	0.76
$\overline{P}_0$	Barrier	23858	25996	5.05	5.30
C_0	Daily electricity cost	\$ 2,767		\$ 0.68	
I_0	Price of mining rig	$\$3.1 \times 10^6$		\$ 1,002	
K_0	Total costs	$\$4.8 \times 10^6$	$\$5.6 \times 10^6$	\$ 1,581	\$ 1,825
T	Maximal mining time	1.87 years		2.65 years	

Table 3 reports the parameter values resulting from these computations along with their counterparts for the baseline calibration. The discounted total costs are lower in the model with exit than in the baseline model. This is not surprising because miners now have a finite horizon. Since costs are smaller, entries happen sooner, which translates into a lower entry barrier. Miners' total costs have two components: the initial purchase of a mining rig and the daily electricity expenditure. We are now able to disentangle them and report their values in the third and fourth lines of Table 3.[32]

It is difficult to get accurate information on how much miners paid for their mining rigs. Yet we can verify that our estimates are in the ballpark of available data. The Antminer S4 ASIC mining rig was released a few days after the beginning of the second period. Thus, taking into account delivery delays, we reckon that most miners were still using the Antminer S3 ASIC early on in the second period. It could perform 0.441 Tera hashes per second and had a power consumption of 340 watts. Depending on the source, the cost of buying an Antminer S3 ASIC varies from \$ 382 to \$ 480.[33] Assuming that miners paid 4 cents per kilowatt hour,[34] they had to spend between 866 and 1080 dollars to buy 1 Tera hash of the Antminer S3 mining rig, and then pay 74 cents of

[32]The estimated value for I_0 depends on r which we set equal to 0.1 as in the baseline model. However, our results are not very sensitive to the choice of r because the obsolescence process is so fast that miners do not operate their machines for a very long time. For example, $r = 0.05$ yields $I_0 = \$1033$ in the second period, while $r = 0.2$ yields $I_0 = \$943$.

[33]The lowest estimate is available on Bitcoin wiki (https://en.bitcoin.it/wiki/Mining_hardware_comparison), whereas the highest is from a reddit forum stating that the Antminer S3 ASIC costs 0.75 bitcoins, which amounted to \$ 480 at that time (https://www.reddit.com/r/Bitcoin/comments/29jni2/i_did_the_maths_for_antminer_s3_potential/).

[34]See https://en.wikipedia.org/wiki/Electricity_pricing for an estimation of the electricity costs. Miners probably manage to pay below the market price for their electricity, but their consump-

electricity per day. These values are indeed very close to the ones resulting from our estimation.

Network electricity consumption.—As Table 3 shows, electricity costs accounted for at most 35% and 37% of entrants' expected discounted costs in the first and second period, respectively.[35] Hence all seignorage income was not spent on electricity, as often argued, but instead largely captured as a monopoly rent by Bitmain, the only manufacturer of ASICs for Bitcoin mining. There is no official source on Bitmain's operating profits because it is not a publicly traded company. The most reliable estimates suggest that Bitmain is indeed a very profitable company that made between 3 and 4 billions US dollars in profits in 2017.[36]

Table 4: Ratio of Operating Costs over Mining Rewards

Configuration	μ	σ^2	a	Operating Costs/Rewards
Calibration	0.19	0.54	0.76	0.37
$a = 0.36$ (Moore's law)	0.19	0.54	0.36	0.50
$\mu = 0$	0	0.54	0.76	0.40
$\sigma^2 = 0.05$	0.19	0.05	0.76	0.34
$\mu = 0$, $a = 0.36$	0	0.45	0.36	0.55

Note: All other parameters are as reported in Table 3 for the model with exit over the second period.

Our model can precisely quantify the share of block rewards that is dissipated in electricity expenditures. Such a statistic is easily recovered from the simulations because they yield the distribution of vintages among incumbents. We focus on averages because the exchange rate (and hence the block reward) are much more volatile than the hashrate. To estimate the share of block rewards devoted to electricity, we simulate many GBM trajectories and take their averages. The initial conditions (namely the initial distribution of hardware vintages and the distance between initial payoffs and the entry barrier) influence the ratio of interest at the beginning of each paths. In order to

tion is also higher than the one needed to run the machines because they also have to cool them down.

[35] Remember that our procedure places an upper bound on expected costs.

[36] By comparison, NVIDIA, a world leader in GPU manufacturing, "only" made 3 billions US dollars in operating profits in 2017. See, among other sources, https://www.cnbc.com/2018/02/23/secretive-chinese-bitcoin-mining-company-may-have-made-as-much-money-as-nvidia-last-year.html

neutralize this dependency, we simulate long trajectories of 10,000 days and discard the first 1,000 observations. Besides the baseline calibration, we perform a series of comparative statics exercises, where we vary the values of one parameter and update $\overline{P}_0$ consequently. Table 4 displays the outcomes of these experiments with the ratio of electricity expenditures over total block rewards in the last column.

Using the calibrated values for the second period, we find that the cross-sectional share of electricity costs over total rewards is 37%. This digit is in line with the results of Table 3. Turning our attention to the comparative statics exercises, we find that the rate of technological progress, a, is again the crucial parameter, with a substantial negative effect on the ratio. We have already shown that the rate of technological progress significantly raises the reflecting barrier $\overline{P}_0$. This increase lowers the share of revenues devoted to operating costs because a higher barrier means less investment in mining hardware and so less machines competing for the same reward. By contrast, the growth rate of the block reward, μ, has a negative impact on $\overline{P}_0$. However, μ also directly increases revenues which raises the denominator of the ratio. Those two effects go in opposite directions so, overall, μ has a weak negative impact on the ratio. Similarly, the volatility of the exchange rate, σ, has a negative effect on the barrier which is compensated by its direct effect on the ratio, as more volatile payoffs are more likely to be well below the barrier.

What lessons can we draw from these experiments regarding the future of Bitcoin's electricity consumption? Sooner or later, the rate of return for holding Bitcoins will have to decline. Similarly, the rate of technological progress will have to slow down and converge, in the best case scenario, to the value predicted by Moore's law. Both adjustments will contribute to increase the share of seignorage income spent on electricity, with long-term ratios above 50%, as displayed in the bottom line of Table 4. Then, should the exchange rate stabilize around $ 40.000 after the next halving, Bitcoin's electricity consumption will reach 1% of the world's consumption. Such levels of pollution are likely to push governments into taking serious steps against the cryptocurrency, thus suggesting that Bitcoin's carbon footprint may eventually place a hard cap on its exchange rate.

1.6 Conclusion

We have shown that the behavior of miners can be approximated using a standard model of industry dynamics with irreversible investment and embodied technological progress. We believe that our findings will be of interest to both economists and Bitcoin practitioners.

For economists, Bitcoin's protocol encodes several features that are rarely observed. Miners mostly face aggregate uncertainty. They operate a technology which exhibits constant returns to scale at the micro-level, and earn revenues that are decreasing in aggregate capacity. All these characteristics make the market for mining a perfect laboratory, all the more so since data are exhaustive, clean and easily available. It is therefore quite reassuring that the canonical model of industry dynamics convincingly replicates the evolution of mining capacity over time.

Our approach also provides a forecasting tool for Bitcoin practitioners willing to invest in mining power. From a practical standpoint, it has three main implications. First, the hashrate of the network is closely related to the exchange rate and, in the event of a significant market crash, the hashrate barely moves in the short run due to the irreversibility of past investments. This is good news for the security of Bitcoin transactions but bad news for their carbon footprint. Second, around two thirds of all seigniorage income is not dissipated in electricity consumption, as often argued, but is instead spent on mining hardware. Third, we expect the energy efficiency of the network to deteriorate as the rate of technological progress inevitably decelerates from the high pace it has experienced so far.

Although our model is fairly accurate in the medium to long run, it sometimes temporarily deviates from the data. These discrepancies arise during periods of high volatility, and so are probably explained by congestion effects and non-linearities in the adjustment cost function. Incorporating these imperfections into our framework is a demanding task since they render the entry barrier state-dependent. Future research should nonetheless strive to estimate such an extension as it would probably improve the model's accuracy during periods of trading frenzy. Finally, another interesting direction for further research would consist in using other cryptocurrencies' markets for mining to calibrate the model, starting with Ethereum as the first obvious candidate.

1.7 Appendix

Proof of Propositions.

Proof of Proposition 1.2 Let $W\left(P_t, \overline{P}_t, A_t\right) \equiv V\left(P_t, t\right) + C_t/r$ denote the value of an entrant net of variable costs as a function of the payoff P_t, the entry barrier $\overline{P}_t$ and the efficiency of the technology A_t. Assumption 4 requires that $dA_t = -aA_t dt$. Assumptions 1 and 3 imply that $dP_t = P_t\left(\alpha dt + \sigma dZ_t\right)$ whenever $P_t < \overline{P}_t$ because Q_t remains constant in that region of the payoff space. Finally, the law-of-motion of the entry barrier $\overline{P}_t$ is endogenous and it is precisely the aim of this proof to show that the market for mining satisfies the equilibrium requirements stated in Definition 1.1 when $\overline{P}_t$ decreases at the rate of technological progress. Thus we conjecture that $\overline{P}_t = \overline{P}_0/A_t$, with $\overline{P}_0$ as in Proposition 1.1, and proceed to show that it is indeed optimal for entrants to wait until $P_t = \overline{P}_t$.

Having specified the law of motion of the three state variables allows us to use Ito's Lemma to derive the Hamilton-Jacobi-Bellman equation satisfied by the value function

$$\begin{aligned} rW\left(P_t, \overline{P}_t, A_t\right) &= P_t + \alpha P_t W_1\left(P_t, \overline{P}_t, A_t\right) - a\overline{P}_t W_2\left(P_t, \overline{P}_t, A_t\right) + aA_t W_3\left(P_t, \overline{P}_t, A_t\right) \\ &\quad + \frac{\sigma^2}{2} P_t^2 W_{11}\left(P_t, \overline{P}_t, A_t\right). \end{aligned}$$

Assume that $\alpha \neq r$.[37] Then the general solution of the Hamilton-Jacobi-Bellman equation reads

$$W\left(P_t, \overline{P}_t, A_t\right) = \frac{P_t}{r-\alpha} + \frac{D_1}{A_t}\left(\frac{P_t}{\overline{P}_t}\right)^{\beta_1} + \frac{D_2}{A_t}\left(\frac{P_t}{\overline{P}_t}\right)^{\beta_2},$$

where D_1 and D_2 are constants whose values will be chosen so as to match some boundary conditions, while β_1 and β_2 are the two roots of the following quadratic equation

$$\mathcal{Q}(\beta) \equiv \frac{\sigma^2}{2}\beta(\beta-1) + (\alpha + a)\beta - a - r = 0.$$

Since $Q(0) = -a - r < 0$ and the coefficient associated to the second order term is strictly positive, we know that one root, β_1 for instance, is strictly positive while the other root, β_2, is strictly negative.

[37] As r tends to α, $\overline{P}_0$ converges to $\left(I_0 + \frac{C_0}{\alpha}\right)\left(\alpha + a + \sigma^2/2\right)$ and $W\left(P_t, \overline{P}_t, A_t\right)$ tends to $\frac{I_0 + \frac{C_0}{\alpha}}{A_t}\left(\frac{P_t}{\overline{P}_t}\right)\left[1 - \log\left(\frac{P_t}{\overline{P}_t}\right)\right]$.

The function W has to satisfy the following three boundary conditions. First, since $\tilde{P}_t = 0$ is an absorbing state, we must have $W(0, \overline{P}_t, A_t) = 0$. This implies that $D_2 = 0$, as otherwise the value function would diverge to either minus or plus infinity when P goes to zero. Second, the left continuity of the value function at the entry threshold $\overline{P}_t$ implies that there can be no arbitrage opportunity solely if the value function is flat at the contact point. This requirement, known as the smooth-pasting condition, is satisfied when $W_1\left(\overline{P}_t, \overline{P}_t, A_t\right) = 0$, i.e. when $D_1 = -\frac{\overline{P}_0}{\beta_1(r-\alpha)}$. Finally, the entry barrier is pinned down by the free entry condition $W\left(\overline{P}_t, \overline{P}_t, A_t\right) = I_t + C_t/r$. This implies $\overline{P}_0 = (I_0 + C_0/r)\frac{(r-\alpha)\beta_1}{\beta_1 - 1}$.[38] Thus we have found a solution which satisfies all the requirements laid-out in Definition 1.1 for the existence of a competitive equilibrium.

Proof of Proposition 1.3 We proceed as in the proof of Proposition 1.2. We assume that $\overline{P}_t = \overline{P}_0/A_t$, for some $\overline{P}_0$, and show that it is indeed optimal for miners to enter the race when $P_t = \overline{P}_t$. The value function of an active miner entered at time τ reads $W\left(P_t, \overline{P}_t, C_\tau\right) = \int_t^{+\infty}\left(\int_0^{\overline{P}_s} \max(x - C_\tau, 0) f^e_{P_s|P_t}(x)dx\right) e^{-r(s-t)}ds$, where $f^e_{P_s|P_t}$ denotes the density of the payoff variable at time s as anticipated by entrants at time t. Under the equilibrium rule, the barrier $(\overline{P}_s)_{s\geq t}$ is deterministic. This is why we do not account for the dependency on the whole future trajectory of the barrier when defining W. We only need to show that $W\left(\overline{P}_t, \overline{P}_t, C_t\right) = W\left(\overline{P}_0, \overline{P}_0, C_0\right)/A_t$ because then the condition $W\left(\overline{P}_t, \overline{P}_t, C_t\right) = I_t = I_0/A_t$ will be met for all t whenever $\overline{P}_0$ is chosen such that $W\left(\overline{P}_0, \overline{P}_0, C_0\right) = I_0$.

According to Assumption 6, potential entrants make their entry decisions based on P_t and $\overline{P}_t$ only. Multiplying the two by the rate of technological progress, this is equivalent to saying that potential entrants make their entry decisions based on A_tP_t and $\overline{P}_0$ only. In this detrended space, the barrier is flat. Hence, under the conjectured rule for entry, the process A_tP_t anticipated by potential entrants is Time-Homogeneous Markov, meaning that $f^e_{A_tP_t|A_sP_s}(y) = f^e_{A_{t-1}P_{t-1}|A_{s-1}P_{s-1}}(y)$. Reinserting this equality into

[38] Alternatively, we could have solved the planner's problem and used the "super contact" condition $W_{11}\left(\overline{P}_t, \overline{P}_t, A_t\right) = 0$.

the definition of W, we find that

$$\begin{aligned}
W\left(\overline{P}_t, \overline{P}_t, C_t\right) &= \int_t^{+\infty} \left(\int_0^{\overline{P}_s} \max\left(x - C_t, 0\right) f^e_{P_s|P_t=\overline{P}_t}(x) dx \right) e^{-r(s-t)} ds \\
&= \int_0^{+\infty} \left(\int_0^{\frac{\overline{P}_u}{A_t}} \max\left(x - C_t, 0\right) f^e_{P_{u+t}|P_t=\overline{P}_t}(x) dx \right) e^{-ru} du \\
&= \int_0^{+\infty} \left(\int_0^{\overline{P}_u} \frac{1}{A_t} \max\left(\frac{y}{A_t} - \frac{C_0}{A_t}, 0\right) f^e_{P_{u+t}|P_t=\overline{P}_t}\left(\frac{y}{A_t}\right) dy \right) e^{-ru} du \\
&= \frac{1}{A_t} \int_0^{+\infty} \left(\int_0^{\overline{P}_u} \max\left(y - C_0, 0\right) f^e_{P_{u+t}A_t|A_tP_t=\overline{P}_0}(y)\, dy \right) e^{-ru} du \\
&= \frac{1}{A_t} \int_0^{+\infty} \left(\int_0^{\overline{P}_u} \max\left(y - C_0, 0\right) f^e_{P_u|P_0=\overline{P}_0}(y)\, dy \right) e^{-ru} du \\
&= \frac{W\left(\overline{P}_0, \overline{P}_0, C_0\right)}{A_t}.
\end{aligned}$$

The second equality follows from $u = s - t$ and replacing $\overline{P}_{u+t}$ by $\overline{P}_u / A_t$. The third and fourth equalities use the change of variable $y = A_t x$. The fifth equality is a direct consequence of the Time-Homogeneous Markov property of $A_t P_t$. The last equality holds by definition and proves that free entry is indeed satisfied when $\overline{P}_t$ decays at the rate of technological progress.

Lemma 1.1. *Let Assumptions 1, 3, 4 and 7 hold true. Then, for all $t > 0$, the density of P_t conditional on the barrier being reached at time $\tau < t$ reads*

$$\begin{aligned}
f^e_{P_t|P_\tau=\overline{P}_\tau}(x) = \left(\frac{1}{x}\right) \Bigg\{ & \left(\frac{1}{\sigma\sqrt{t}}\right) \phi\left(\frac{\log(\overline{P}_\tau) - \log(x) + \left(\alpha - \frac{\sigma^2}{2}\right) t}{\sigma\sqrt{t}} \right) \\
& + \exp\left[\left(\log(\overline{P}_\tau) - \log(x) - at\right) \left(1 - 2\left(\frac{a+\alpha}{\sigma^2}\right)\right) \right] \\
& \times \Bigg[\left(2\left(\frac{a+\alpha}{\sigma^2}\right) - 1\right) \Phi\left(\frac{\log(x) - \log(\overline{P}_\tau) + \left(2a + \alpha - \frac{\sigma^2}{2}\right) t}{\sigma\sqrt{t}} \right) \\
& + \left(\frac{1}{\sigma\sqrt{t}}\right) \phi\left(\frac{\log(x) - \log(\overline{P}_\tau) + \left(2a + \alpha - \frac{\sigma^2}{2}\right) t}{\sigma\sqrt{t}} \right) \Bigg] \Bigg\} \mathbb{1}_{]0,\overline{P}_t]}(x),
\end{aligned}$$

where ϕ and Φ are the density and the cumulative distribution function of the standard normal distribution, respectively.

Proof of lemma 1.1 Since Assumption 7 implies Assumption 6, Proposition 1.3 applies and we know that there exists a $\overline{P}_0$ such that $(P_t, \overline{P}_t = \overline{P}_0 / A_t)$ is an industry equi-

librium. Moreover, Assumption 7 also implies that the anticipated P_t follows a GBM when $P_t < \overline{P}_t$ because the hashrate Q_t remains constant. Hence the anticipated P_t follows a GBM reflected at $\overline{P}_0/A_t$. The density of a positive Brownian motion reflected at 0 and which starts at 0 is given in Harrison (2013). We now show that it can be applied to the logarithm of P.

Without loss of generality, we can set the hitting time $\tau = 0$. Then $R_0/Q_0 = \overline{P}_0$ because we are looking for a density conditional on $P_0 = \overline{P}_0$. Hence the hashrate Q_t is given by

$$Q_t = \sup_{0 \leq s \leq t} A_s R_s / \overline{P}_0.$$

Replacing this expression into the decomposition of $A_t P_t$, we find that

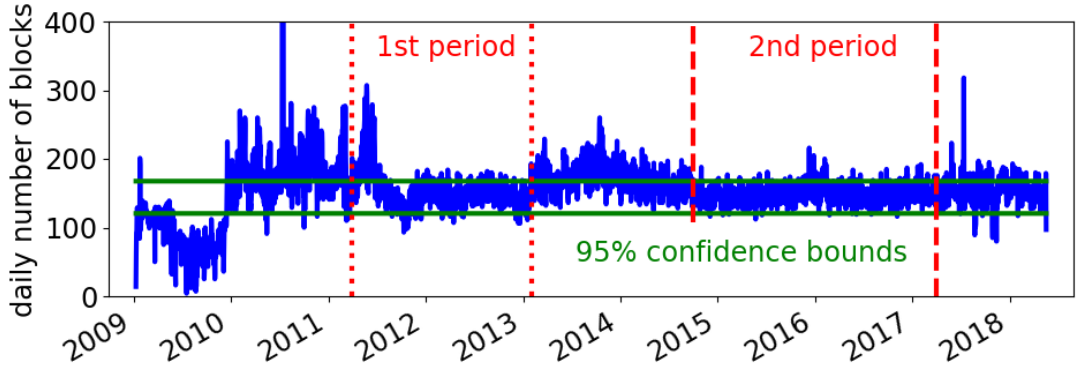
$$\begin{aligned}
\log(A_t P_t) &= \log(A_t R_t) - \log(Q_t) \\
&= \log(A_t R_t) - \sup_{0 \leq s \leq t} \log(A_s R_s) + \log(\overline{P}_0) \\
&= \log(\overline{P}_0) - \left[-\log(A_t R_t) - \inf_{0 \leq s \leq t} (-\log(A_s R_s)) \right] \\
&= \log(\overline{P}_0) - Z_t,
\end{aligned}$$

where Z_t follows a positive Brownian motion with parameters $(\sigma^2/2 - a - \alpha, \sigma)$, reflected at 0 and with initial condition $Z_0 = 0$. We know from Harrison (2013) that, for all $x \geq 0$, $\Pr(Z_t \leq x) = \Phi\left(\frac{x - \left(\frac{\sigma^2}{2} - a - \alpha\right)t}{\sigma\sqrt{t}}\right) - e^{\frac{2\left(\frac{\sigma^2}{2} - a - \alpha\right)x}{\sigma^2}} \Phi\left(\frac{-x - \left(\frac{\sigma^2}{2} - a - \alpha\right)t}{\sigma\sqrt{t}}\right)$. Straightforward differentiation of this expression yields the solution for f^e.

Test of Assumption 1.

According to Assumption 1, finding a block would always take 10 minutes on average so that the daily number of blocks found would not be statistically different from 144. Figure 13 plots the daily number of blocks found along with the two 95% confidence bounds. For our periods of interest, the results are satisfying except for the beginning of the first period. According to this graph, it is sensible not to consider the period in between when ASICs were introduced. Then technological progress was so fast that the hashrate significantly exceeded the target of one block every ten minutes.

Figure 13: Number of blocks found per day

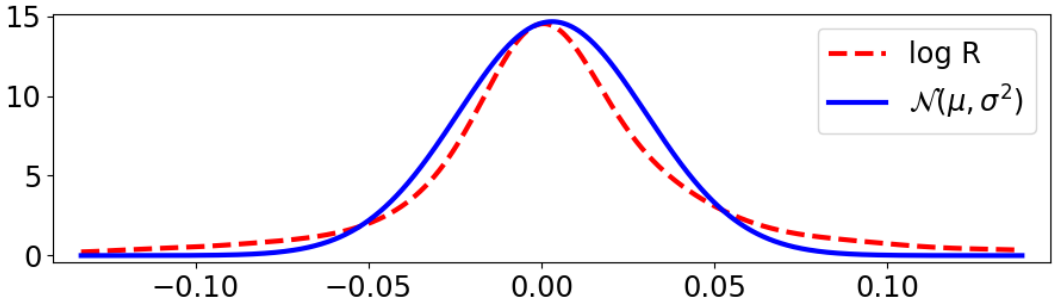


Note: the number of blocks fund per day has been retrieved from coindesk.com.

Test of Assumption 3.

The GBM assumption implies that the log returns of block reward should be both independent and normally distributed. We first show that the distribution of log returns can be well approximated using a normal distribution. Figure 14 compares the non-parametrically estimated density of log returns with their normal density estimated under the GBM assumption. For our parametric estimation, we exclude some extreme events by discarding the 5% most extreme returns on each side. This procedure yields densities that are quite close to each other.

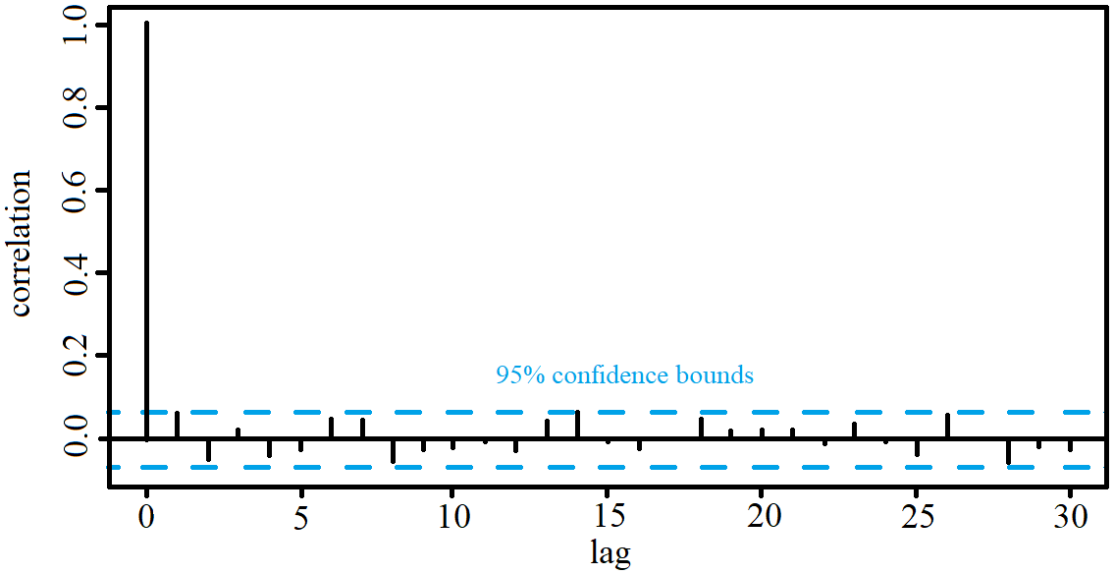
Figure 14: Normality of returns



As for the independence property, Figure 15 shows that log returns are not linearly autocorrelated. We obtain similar results composing the log returns with many other functions. However, statistical tests indicate that the variance of the exchange rate does not remain constant over time, and goes instead through periods of high and

low volatility. Although the issue is strongly alleviated by our division of the sample into two subperiods, it suggests that a more realistic specification should allow the variance coefficient σ to vary over time. We leave this extension to further research because it makes the barrier state dependent and thus greatly complicates the characterization of the equilibrium.

Figure 15: Independence of returns

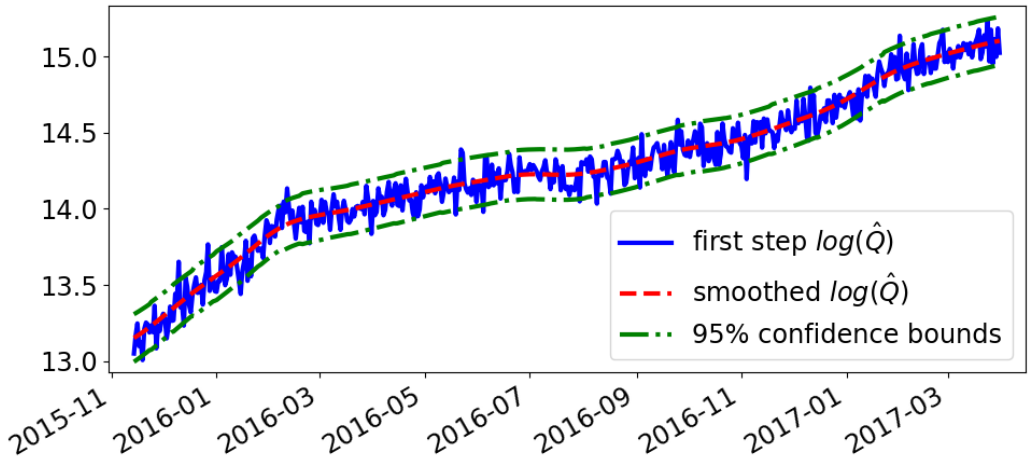


Note: his autocorrelogram has been obtained using the second period log returns.

Estimation of Q.

$(Q_t)_{t\geq 0}$ is not observable but can be estimated using a two-step procedure. First, for each day t, let $\hat{Q}_t \equiv N_t/\tilde{\Pi}_t$, where N_t is the number of blocks found for day t and $\tilde{\Pi}_t$ is the probability to find a valid block with a single hash. Both are directly observable in the blockchain. Since $N_t \sim \mathcal{B}i\left(Q_t, \tilde{\Pi}_t\right)$, $\hat{Q}_t$ is a very natural estimator of the daily hashrate. This estimator is non biased and it can easily be shown that it is asymptotically equivalent to the maximum likelihood estimator. Given that there is a lot of variation across daily estimates, we smooth this new time series using a local linear regression. Figure 16 shows we are not losing much information performing a local linear regression over $\hat{Q}$.

Figure 16: Estimation of Q



The two green curves are confidence bounds for the first step estimation if the true $(log\,(Q)_t)_{t\geq 0}$ were the red curve (the second step estimate). If the erratic variations of the first step estimation captured not only the first step estimation variance but also some real variations of the hashrate not captured by the second step estimation, then its variance should be bigger than the one resulting from the first step estimation error only. Thus it should cross the green bounds much more often than 5% of the time, which does not happen in our data. For the sake of clarity, we do not show the whole series but the test works very well for the whole period.

2 How to make the Bitcoin network more environmentally friendly

Abstract

I start modeling very simply the interactions between the different Bitcoin stakeholders to emphasize the inefficiency in the current Bitcoin protocol yielding to a terrible waste of money and electricity. I then show that the first best protocol is unfortunately unreachable due to an incentive-compatibility constraint. I pin down the second best protocol and show that it is preferable to the current protocol as soon as the yearly technical progress rate as regard mining hardware is above 10%.

2.1 Introduction

Bitcoin was created in the end of 2008 / beginning of 2009 by Satoshi Nakamoto (see Nakamoto (2008)). It is the first viable currency which needs not rely on banks (central or commercial) to work. It was first designed to ease online commerce. On top of the absence of intermediary, which means lower costs, sellers need not be wary of their customers and hassle them with many personal details because bitcoin payments are irreversible. But Bitcoin can provide many more financial services than mere monetary transfers. The use of smart contracts (originally invented by Szabo (1996)) enables the sender of a payment to specify complex conditions that the recipient needs to satisfy to redeem the coins. Nowdays, many second-layer protocols, which enable users to perform a bunch of fancy operations, rely on Bitcoin, making it the backbone of an ever-growing ecosystem of new technologies. To cite a few, the lightning network (Poon and Thaddeus (2016)) enables users to send free and instant payments, Rootstock (Lerner (2015)) extends the possibilities offered by smart contracts, coloured coins Assia et al. (2013) allow users to exchange real assets using Bitcoin and Tether is a cryptocurrency pegged to the US dollar.

Miners are the ones who make the Bitcoin network secure by providing it with computing power. Together, they solve difficult cryptographic puzzles. If an attacker wants to cancel one of her payments, she needs to solve those cryptographic puzzles faster

than all the other miners together. This way of securing transactions is called "Proof-of-Work". The more computing power (called the "hashrate") miners deploy, the more secure Bitcoin transactions are. Miners are rewarded for their work. Every ten minutes on average, one randomly chosen miner earns a predetermined reward. The probability for a miner to win the reward is proportional to her computing power. Prat and Walter (2018) model the market for mining and show that the total hashrate of the network can be very well predicted with the reward scheme. The reward increases with the Bitcoin / US dollar exchange rate. Free entry on the market for mining implies that after an increase in the Bitcoin / US dollar exchange rate, more miners start mining and so more electricity is spent on this activity.

If the total hashrate is high enough, everybody reckons that no single entity can outperform honest miners and so everybody sees bitcoin transaction as absolutely irreversible. A higher hashrate would then be sub-optimal. Bitcoin would not be deemed safer but more electricity would be wasted on useless computations. I believe that by the time those lines are written - February 2018 - the hashrate is much too high, as regard the previous criterion. As a result: miners' electricity consumption is tremendous: Bitcoin consumes as much electricity as Morocco or Bulgaria! Of course, Bitcoin is deemed extremely safe, but this was already the case a couple years ago, when the electricity consumption was not a tenth of what it is today!

Electricity being consumed on securing transactions is inherent to the Proof-of-Work algorithm. Other algorithms, much less electricity-dependent are available. Ppcoin, also known as Peercoin, (see King and Nadal (2012)) uses the Proof-of-Stake algorithm, where coin age (the number of coins an entity holds multiplied by the age of those coins) is substituted to computing power. No miners are needed to secure Peercoin. Sunny King also introduced Primecoin (see King (2013)), which relies on Proof-of-Work, but where the computations performed are useful (as opposed to Bitcoin) since they consist in finding sequences of prime numbers, which can be of great interest for mathematicians. IOTA resorts to a directed acyclic graph instead of a blockchain to secure transactions. See Popov (2017). This approach is also much environmental-friendlier than Proof-of-Work.

It seems very unlikely to see Bitcoin switch to Proof-of-Stake, for instance, in the

near future. This would represent a substantial change in the protocol and this idea is not in the pipes. I argue that even within the current Proof-of-Work setting, it is easy to dramatically lower miners' electricity consumption without giving up on security simply by lowering miners' rewards. By the time those lines are written, every ten minutes, one of the miners wins about 112,000 US dollars. In January 2018, this reward briefly reached 300,000 US dollars! No surprise it creates a huge incentive for miners to mine. Why is the reward so high? First, because bitcoins are very expensive. The role Bitcoin plays within the fintech industry shows that it has a fundamental value. Yet, many people argue that Bitcoin is a bubble ready to burst. The point of this article is not to answer this question. Second, Satoshi Nakamoto designed the whole reward scheme in the beginning of 2009 and the community has failed to modify it since then. If Nakamoto's scheme made much sense at the beginning, it is not adapted to the current Bitcoin / US dollar exchange rate any more. Nakamoto could obviously not predict in 2009 the entire path the exchange rate would follow.

Modifying the Bitcoin protocol to lower miners' reward is not an easy task since, due to their central role, miners have the power to prevent any protocol change. Protocol change proposals must then be made incentive-compatible with respect to the miners. For that matter, I suggest to first give miners a substantial bounty before lowering their rewards, in order to compensate them for the drop in their incomes. I show that in most cases such a scheme can be made incentive-compatible a lead to much electricity being saved.

The rest of the article is organized as follows. Section 2 briefly explains how Bitcoin works and highlights the role of miners and the energy waste caused by their ill-designed reward scheme. In section 3, I model the interactions between the different Bitcoin stakeholders and compare the current Bitcoin protocol with the unreachable first best one. In section 4, I find the second best protocol and show to what extent it is better than the current one. Section 5 concludes.

2.2 Bitcoin and the miners

This section describes the role miners play in the Bitcoin protocol. The goal of this article is not to explain how Bitcoin works since this task has already been successfully

fulfilled many times. I refer the interested reader to Nakamoto (2008) and Antonopoulos (2014). I only explain what is necessary to know to understand the rest of the article.

Bitcoin needs not rely on banks. This property is known as "decentralization" since no single central authority controls the currency. If such a design supposedly raises security, it creates one main problem: how to prevent a user from spending the same coin twice? A user can try spending a coin twice by broadcasting on the network two conflictual transactions simultaneously. Since the two transactions are conflictual, one of the two will eventually be discarded. The double-spend attempt succeeds if the recipient of the transaction which will be discarded accepts the payment and delivers the corresponding good / service. For the system to work efficiently, a recipient of a transaction must know very quickly whether the transaction is fraudulent or not. The solution to this problem Satoshi Nakamoto proposed relies on the Blockchain technology and on the Proof-of-Work mechanism.

Transactions are not added to the Blockchain (the full history of all transactions) one by one, they are added by blocks. Producing a valid block is, on purpose, very difficult. Each block possesses a header, which contains an arbitrary integer (called a nonce) and a statistic which sums up all the transactions of the block, the time the block was assembled and the header of the previous block. Finding a valid Proof-of-Work boils down to finding a nonce such that $h(header) \leq t$, for some threshold t, where h is the sha-256 hash function applied twice. As a hash function, h is numerically non invertible. On top of that, knowing $h(n)$ for some $n \in \mathbb{N}$ yields no information on $h(m)$, for all $m \neq n$. As a result, there exists no smart algorithm to find a valid Proof-of-Work. The only method available is to increase the nonce and hash the header until the condition $h(header) \leq t$ is satisfied. This problem can be made arbitrarily difficult by lowering the threshold t. Trying to find valid blocks is called mining, and individuals, or entities, who use their computing power to accomplish this task are called miners.

How do they make transactions secure? Since the header of a block contains the hash of the header of the previous block, blocks are cryptographically chained. It means that if a block is tampered with, not only its hash will not be under the thresh-

old anymore, but all the blocks later in the chain will also become invalid. When the network is aware of two competitive chains, all nodes select the longest one (in terms of the number of blocks). Thus, to be able to erase a transaction from the blockchain, an attacker would have to build himself a blockchain longer than the main one. She has to recompute the block in question as well as all the blocks found later... and this before one of the honest miners finds one more block. So if honest miners account for a large majority of the computing power, a transaction part of a block buried under a couple other blocks is deemed irreversible.

Mining is a costly activity. It requires acquiring the right hardware and providing it with electricity. Therefore, mining must be rewarded. When a miner creates a valid block, she earns a fixed amount of newly created coins and the sum of all fees granted by the transactions included in the block. Transactions fees are freely chosen by users. Nakamoto's idea was the following: in the long term miners must be rewarded only with transactions fees, so users pay for the service they use. But in the short term, before Bitcoin is famous enough to generate enough transaction fees, to keep the system secure miners must also earn the newly created coins. This way of remunerating miners also solves the money creation problem. Nakamoto chose the following design. At the beginning, each block yields 50 new bitcoins to the miner who finds it. Every 210,000 blocks (around 4 years), this reward is divided by two, so that the reward scheme slowly converges towards the long term solution. It is easy to see that the monetary base will converge to 21 millions. Since the inception of Bitcoin, this reward design has never been adapted, although the number of new coins awarded to miners is nothing more than a parameter which can be freely chosen. It is by no means a fundamental feature of Bitcoin and it is not even discussed in Nakamoto (2008). Nowdays, Bitcoin is probably famous enough to implement, or at least move close to, Nakamoto's long term solution. Yet, miners still receive 12.5 new bitcoins per block.

Since they are the ones who create blocks, miners have a big say on the rules which define the Bitcoin protocol. Imagine that a substantial number (in terms of hashrate) of miners do not like the new protocol and keep following the old protocol. Those miners will reject blocks created by the miners who follow the new rules and they will build their own blocks, that the miners who follow the new rules will reject. The blockchain

will split in two, which can be detrimental due to the network effect and create confusion among users. Only if the new protocol is a subset of the old one (every block or transaction valid under the set of new rules is still valid under the set of old rules) and a majority of miners follow the new rules can a change of protocol be successful without almost all the miners agreeing, due to the longest chain rule. The protocol change I advertise is unfortunately not a subset of the current protocol. I need all miners to agree on it so none of them must be worse-off.

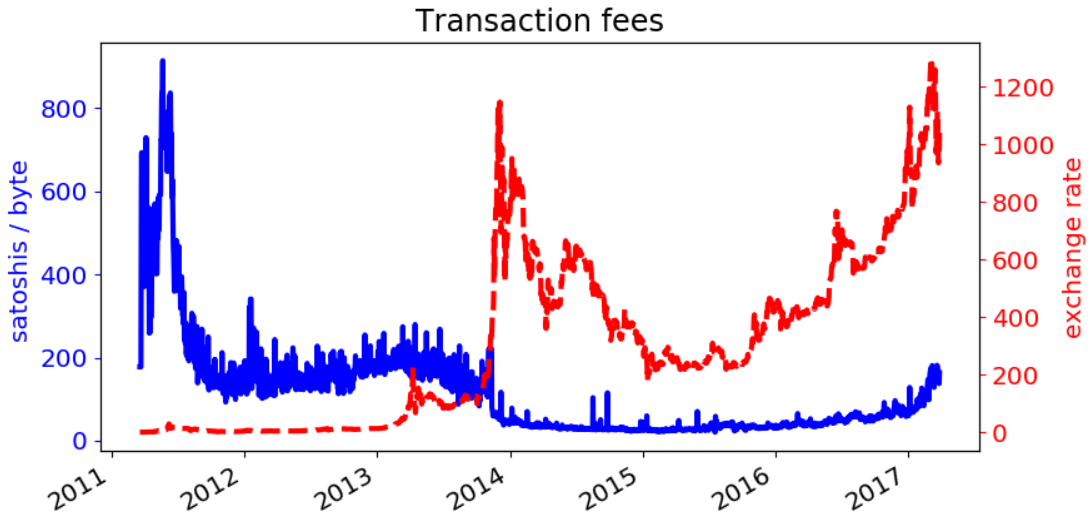
2.3 The model

I start modeling the interactions between the bitcoin holders, the users and the miners, to underline the inefficiency of the current situation. Of course, in reality users and holders are, at least partially the same people since one needs to own bitcoins in order to use them. Yet, in this model, it is easier to consider them separately. The model uses discrete time. Let $t = 0$ denote the time period the protocol change I advocate is announced. I model the interactions between the different Bitcoin stakeholders from this day on only. What happened before does not matter here. The model is extremely simple and often far from reality. Yet, the way agents behave and variables of interest evolve has been changing fast. I argue that the gap between the model and reality is a bit narrower for the most recent period. So the model should be assessed on this period only.

Market capitalization and exchange rate Like Prat and Walter (2018), (from now on abbreviated "PW") I take the ₿/ $ exchange rate as exogenous. The extreme volatility of this exchange rate for sure influences miners' behavior. The two previous authors took this phenomenon into account and modeled miners' reward as a geometric brownian motion (GBM). Proceeding so was a basic requirement for them since the main point of their article was to provide a model which uses the exchange rate to reproduce miners' observed investment behavior. In this article, the exchange rate is deterministic. The first reason to motivate this choice is that the exchange rate plays a much less central role here than in PW. The second reason is that here the reward miners earn is decomposed into the fees and the newly minted bitcoins. Assuming that the sum of the two follows a GMB implies that the fees alone follow a very not natural

process, which may complicate a lot the analysis since the fees play a more central role in this article then in PW. Last, the mechanism I want to study is pretty simple and so I prefer focusing on clarity and simplicity instead of fidelity towards reality. I denote K, the market capitalization of Bitcoin (the value of all the bitcoins). It is assumed to grow at a rate α. That is to say, we have $K_t = K_0 \times (1+\alpha)^t$, where K_0 is exogenous. Let B_t denote the number of new bitcoins minted at date t. The ₿/ $ exchange rate, which I denote R_t, is equal to the market capitalization divided by the number of bitcoins in circulation. We have $R_t = \frac{K_t}{\mathcal{B}_0 + \sum_{s=1}^{t} B_s}$, where $\mathcal{B}_0$ is the total number of bitcoins at time 0.

Transaction fees Transaction fees are freely chosen by users. They are used to clear the block space market. Indeed, blocks are limited to 1 MB. When more than 1 MB of transaction data are broadcast to the network every ten minutes, all transactions cannot be included into blocks. Miners select the ones which pay the highest fees. For simplicity, the $ value of the total transaction fees per period is constant over time in the model. This is definitely a simplifying assumption but I reckon that the gap between this assumption and reality, not too wide, does not affect much miners' behavior. First, note that users have in mind the ₿/ $ exchange rate when they pick transaction fees. If the ₿/ $ exchange rate increases, everything else equal, the amount of fees, paid in bitcoins, decreases. This fact is illustrated on the following figure, which plots the average number of satoshis (10^{-8} bitcoins) given as fee per byte of transaction data and the ₿/ $ exchange rate.



On the long term, the exchange rate goes up and the amount of bitcoins given as fees per byte goes down. This is particularly visible in 2013. The rise in fees in 2017 is due to congestion: at that time the size of transaction data broadcast every ten minutes was getting close to the upper limit. But what matters here is not the fees per byte of transaction data but the total amount of fees per time period. This last quantity of course depends on the number of transaction broadcast per time period. Since the beginning of 2017, most blocks are very close to 1 MB and the demand for block space is unlikely to go down in the near future. Second, in reality, the level of transaction fees is very volatile even within a single day, depending on the instant demand for block space and how lucky miners are finding blocks (instant block space offer). But due to the irreversibility of investments in mining hardware, the network hashrate enjoys a certain inertia and does not respond to the very short term volatility in revenues induced by transaction fees. Miners would behave in a very similar way, were the level of transaction fees flat, at its average.

The bitcoin holders Bitcoin holders maximize their wealths. In this model, they all have symmetric roles. Maximizing separately the wealth of each bitcoin holder is equivalent to maximizing the sum of the wealths. For the sake of simplicity, I assume that a single agent owns all the bitcoins. This agent is not financially constrained: she is a long term investor who does not have to sell some of her bitcoins at an early date. She wants to maximize her wealth at a far-away date T. The exact date T does not matter much. With the protocol change I propose, the agent will be a bit worse-off than with the current protocol in the short term but much better-off quite quickly. The far-away date T only highlights the fact that the short term does not matter here. At every period, the holder buys all the newly minted bitcoins. Her wealth at date t reads $W_t = K_t - \sum_{s=0}^{t} R_s B_s$. The value of her wealth does not matter. What matters is the difference between the current protocol and the one we propose. That is why it is not necessary to subtract from her wealth all the bitcoins mined before before the date $t = 0$.

The miners The mining industry enjoys constant returns to scale (two pieces of mining hardware will generate valid blocks twice as often as one piece of hardware will), so for a given miner, it does not matter who (herself or somebody else) owns the other

machines. Consequently, I assume that at each period t, a new miner has the opportunity to enter the race and buys a quantity $q_t \in \mathbb{R}_+$ of hashpower. A machine which delivers one unit of hashpower costs $\tilde{I}_t$, at time t. To operate this machine, the miner must pay C_t per unit of hashpower for electricity at every period. Those pieces of hardware are specifically designed to mine Bitcoin. As a result, they are incredibly fast but cannot be used for any other purpose. Thus those machines cannot be resold. Buying hashpower constitutes an irreversible investment. Mining solo is very risky since one may well never find any valid block. To smooth their revenues, miners are all parts of a few big mining pools, which share revenues according to individual hashpower. In March 2018, the 10 biggest mining pools represent around 95 percent of the network hashrate. After a couple weeks, there is no more uncertainty left. Miners' revenues are exactly proportional to their computing powers.

Let Q_t denote the network hashrate at time t. At each period s, a miner who has entered the race at time t earns $q_t \left(R_s B_s + F\right)/Q_s$ and pays $q_t C_t$ for electricity. The mining hardware enjoys a constant rate of technological progress, a. That is to say, we have $\tilde{I}_t = \tilde{I}_0/(1+a)^t$ and $C_t = C_0/(1+a)^t$. I assume that miners cannot stop mining.[39] Let r denote the discount factor. The profit of a miner who enters the race at date t reads:

$$\begin{aligned}
\Pi_t &= q_t \left(\sum_{s=t}^{+\infty} \frac{\left(\frac{R_s B_s + F}{Q_s} - C_t\right)}{(1+r)^{s-t}} - \tilde{I}_t \right) \\
&= q_t \left(\sum_{s=t}^{+\infty} \frac{R_s B_s + F}{Q_s (1+r)^{s-t}} - \left(\frac{C_t(1+r)}{r} + \tilde{I}_t \right) \right) \\
&= q_t \left(\sum_{s=t}^{+\infty} \frac{R_s B_s + F}{Q_s (1+r)^{s-t}} - I_t \right),
\end{aligned}$$

where $I_t = \frac{C_t(1+r)}{r} + \tilde{I}_t$ is the total discounted cost paid by a miner. We still have $I_t = I_0/(1+a)^t$ so this last equality shows that assuming that miners are not allowed to stop mining (even when they lose money) is equivalent to assuming that electricity is free. In both cases, the hashrate can never decrease and we have $Q_t = Q_0 + \sum_{s=1}^{t} q_s$. Free entry translates into

$$\Pi_t = 0, \text{ for all } t \geq 0 \tag{11}$$

[39] This assumption is discussed in detail in PW. It simplifies the model a lot and it does not affect the network hashrate much.

An equilibrium on the market for mining is a Nash equilibrium of the game where miner t's strategy is $q_t \geq 0$. If Q_{-1} is known, knowing the path $\{q_t,\ t \geq 0\}$ is equivalent to knowing the path $\{Q_t\ , t \geq 0\}$. I use the latter in the formal definition of an equilibrium.

First, let denote the whole network earnings $E_t \equiv R_t B_t + F$. And from now on, the symbol "*" as an exponent is used to denote variables at equilibrium.

Definition 2.1 (Equilibrium on the market for mining)**.** *An equilibrium on the market for mining for an earnings path* $\{E_t,\ t \geq 0\}$ *and an initial hashrate* Q_{-1} *is an increasing hashrate path* $\{Q_t^*,\ t \geq 0\}$ *such that equation 11 holds and* $q_t^* = 0$ *implies* $\sum_{s=t}^{+\infty} \frac{E_s}{Q_s^*(1+r)^{s-t}} - I_t \leq 0$.

The users The users send / receive bitcoins and pay transaction fees, F per period, to miners. Here again, all users are symmetric so it is simpler to consider a single user. Of course this user likes safe transactions. Let $\overline{Q}_t$ denote the threshold hashrate at time t, above which mined transactions are considered totally irreversible. The utility of the user at time t reads $U_t = \beta \min\left(\frac{Q_t}{\overline{Q}_t}, 1\right) - F$, for some positive β. The "total cost of all the running machines", for a hashrate level Q_t is $I_t Q_t$. This quantity is a qualitative indicator of how much it would cost an attacker to try double-spending coins. I define $\overline{Q}_t$ as the minimum hashrate such that the total cost of all running machines is higher than a certain security level S, so $\overline{Q}_t I_t = S$. We see that $\overline{Q}_t$ increases at the technical progress rate. I discuss later what would be the value for S.

Note that I will not have to aggregate a level of wealth, a profit and a utility since the protocol change I advocate Pareto-dominates the current one in the model.

First best protocol and inefficiency of the current one. I now state that transaction fees are high enough to sustain a hashrate Q_t higher that $\overline{Q}_t$, even if they are miners' only income source.

Assumption 8. $F \geq S\frac{a+r+ar}{(1+a)(1+r)}$

Lemma 2.1. *Under assumption 8, if miners earn transaction fees only (that is to say, we have* $E_t = F$*, fo all* $t \geq 0$*), and if the initial hashrate* Q_{-1} *is low enough* ($Q_{-1} \leq$

$\frac{F}{I_0}\left(\frac{(1+a)(1+r)}{a+r+ar}\right)$*), there exists an equilibrium on the market for mining for which* $Q_t^* \geq \overline{Q}_t$*, for all* $t \geq 0$.

Proof. Let consider the hashrate path $\left\{Q_t = \frac{F}{I_t}\left(\frac{(1+a)(1+r)}{a+r+ar}\right),\ t \geq 0\right\}$. This path is strictly increasing. At each date, the miner who has the opportunity to enter the race does it. A simple computation shows that for all t, $\sum_{s=t}^{+\infty} \frac{R_s B_s + F}{Q_s(1+r)^{s-t}} - I_t = 0$. We have indeed an equilibrium. Moreover, assumption 8 yields $Q_t \geq \frac{S}{I_t} = \overline{Q}_t$. □

Is assumption 8 likely to be satisfied in reality? This, of course, depends on the value of S, which is hard to know exactly. The following table gives approximate values for $Q_t I_t$, for each year since 2011.

Table 5: Approximate security level $(Q_t \times I_t)$ by year

time	security level in $M
01/2011	.15
01/2012	.4
01/2013	1
01/2014	40
01/2015	200
01/2016	550
01/2017	890
01/2018	2150

We see that the cost of spending coins twice has always increased, along with miner's earnings. For already many years, Bitcoin has had the reputation of being extremely safe. To be honest, until 2015, most articles in mainstream economic journals rather highlight the different frauds and bitcoin thefts. However, those are linked to the (bad) way some bitcoins are stored. They are totally independent from the Bitcoin protocol. As a matter of fact, there has apparently never been any successful double-spending attempt. If I consider (quite conservatively) that S is the security level prevailing in the beginning of 2015, $(S = 200)$ we see that the current hashrate can be divided by ten without any risk. With such a value for S, assumption 8 is by far satisfied, considering the high level of transaction fees miners enjoy in the beginning of 2018, because those account for much more than 10 percent of miners 'rewards.

On top of that, were assumption 8 not quite satisfied, the rest of this article would still make sense. Instead of setting the number of new bitcoins created per block to 0, as is suggested a bit further, one could set it to a strictly positive value, well below its current value but high enough in order to maintain security. Even under the current Bitcoin protocol, the number of new bitcoins created per block will converge to 0. So if assumption 8 is not satisfied, the current protocol may not be viable in the long term, were the minimum fee level not raised.

The equilibrium I highlighted is not the only one. For example, one can find an equilibrium where new miners enter the market every other day only. I will still consider only the equilibrium of lemma 2.1 since it is the most natural one and the only Markov-perfect one. Besides, for other equilibria, the hashrate path would undulate around the equilibrium hashrate path I gave. Provided those undulations are small enough (which is the case for the equilibrium where miners invest every other day), there is no significant qualitative difference between all the obtained equilibrium hashrate paths.

Let $P_t = \frac{R_t B_t + F}{Q_t}$, be miners' per period payoffs.

Assertion 2.2 (Equilibrium selection)**.** *Whenever it is available (whenever it does not require the hashrate to decrease), miners coordinate on the equilibrium where* $P_t = P_0/(1+a)^t$*, for some* P_0*.*

I now compare the current protocol with the one I advocate. A Bitcoin protocol is a set of rules defining precisely what is a valid transaction and what is a valid block. In the model, the only rule that can be changed is how many new bitcoins are miners allowed to give themselves per block.

Definition 2.3 (Protocol)**.** *A protocol is a path* $\{B_t \geq 0,\ t \geq 0\}$ *for the number of new bitcoins miners earn per block.*

So far, I consider two different protocols: the current protocol and the "utopic protocol". To describe them, let assume a time period is a day.

Current protocol. We have $B_t = 1800$[40] until the next reward halving, then $B_t = 900$ for four years, then $B_t = 450$, and so on. At equilibrium, the hashrate is always above $\overline{Q}$ thanks to lemma 2.1 and miners' earnings are always bigger (much bigger at the beginning) than F. So the user's utility is $U_t = \beta - F$. The holder's wealth at time t is $W_t = K_t - \sum_{s=0}^{t} B_s R_s$. And miners always earn 0, by definition of an equilibrium.

Utopic protocol. It is defined by $B_t = 0$ for all $t \geq 0$. If this protocol were anticipated for long enough, Q_{-1} would be low enough, as in lemma 2.1 and the equilibrium hashrate would be the one displayed in lemma 2.1. The user's utility would still be $U_t = \beta - F$. The holder's wealth would be $\tilde{W}_t = K_t > W_t$, for all $t > 0$. Miners would still earn 0.

The utopic protocol Pareto-dominates the current protocol. It is easy to see that the utopic protocol is actually the Pareto-optimum. The current protocol is very inefficient. At every period, the holder must pay a substantial "tax" $B_t R_t$, which does not increase the utility / profit of anybody else and is only used to buy mining hardware and burn electricity.

Unfortunately, the today's hashrate, Q_{-1} is already way too high. If we were to switch tomorrow, for all equilibria that could arise, active miners would make losses. This is particularly problematic since miners have the power to block any protocol change. This is why I design a third protocol, the "feasible protocol", which takes into account this incentive-compatibility constraint. I now explain how to write this constraint easily, although miners all start mining at different times. From now on, the symbols "CP" and "FP" as exponents are used to denote variables in the current protocol and feasible protocol respectively.

At time t, the fraction of their incomes active miners have not earned yet is $q_t \sum_{s=t}^{+\infty} \frac{R_s B_s + F}{Q_s (1+r)^{s-t}}$, for all of them. This also holds for the miner who has the opportunity to enter at time t, once the cost I_t has been paid. Before any protocol change is anticipated, the market for mining is at the current protocol equilibrium, for which $\sum_{s=t}^{+\infty} \frac{R_s B_s + F}{Q_s (1+r)^{s-t}} = I_t$, for all t. As a result, we must only check whether the miner who has the opportunity to enter the market at time $t = 0$, when the future protocol change is announced, is indeed willing to do so. If so, no miner will be worse-off.

[40] This corresponds to 12.5 new bitcoins per block and 144 blocks per day on average.

The incentive-compatibility constraint reads:

$$\sum_{t=0}^{+\infty} \frac{R_t B_t + F}{Q_t^{*FP}(1+r)^t} = I_0 \tag{12}$$

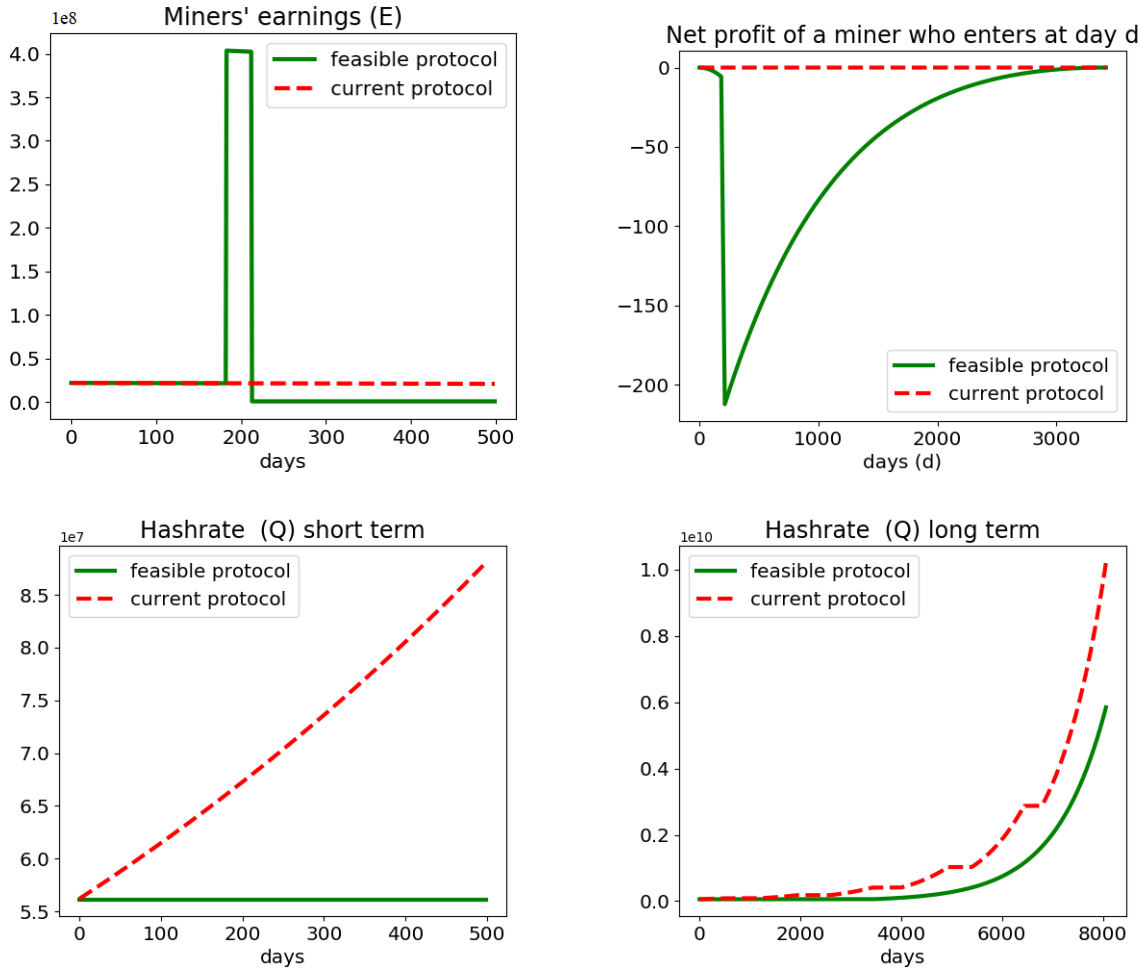
Feasible protocol. We have $B_t = \begin{cases} 1800 & \text{for} \quad 0 \leq t < t_1 \\ \overline{B} >> 1800 & \text{for} \quad t_1 \leq t < t_2 \\ 0 & \text{for} \quad t \geq t_2 \end{cases}$.

I assume that the effective protocol change happens only at $t = t_1$ but that the idea to change the protocol is made public at time $t = 0$ already. Miners anticipate the new reward scheme from $t = 0$ onward. In reality, time is always needed between the moment when a protocol change is announced and the moment it becomes effective, because all users must upgrade their softwares. I fix $t_1 = 183$ (6 months). There are still two parameters: t_2 and $\overline{B}$. These should be determined so as to maximize the holder's wealth under the incentive-compatibility constraint. The theoretical solution for t_2 is easy to get. The holder must transfer some wealth to miners and we are trying to minimize this transfer. Due to the presence of the discount factor, the sooner the transfer is done, the more miners will value it, the smaller it can be. So theoretically, t_2 should be as close as possible to t_1. Yet, in reality it would not make sense to have $t_2 = t_1 + 1$. Even though miners mine in pools, there remains a bit of uncertainty as regards when blocks are found. This uncertainty can become problematic if the bounty period is very short. I thus pick $t_2 = t_1 + 30$. Now, as long as the incentive-compatibility constraint is satisfied, the holder's wealth is clearly decreasing in $\overline{B}$. I proceed numerically to find $\widehat{B}$: the lowest $\overline{B}$ which satisfies the constraint.

2.4 Numerical analysis

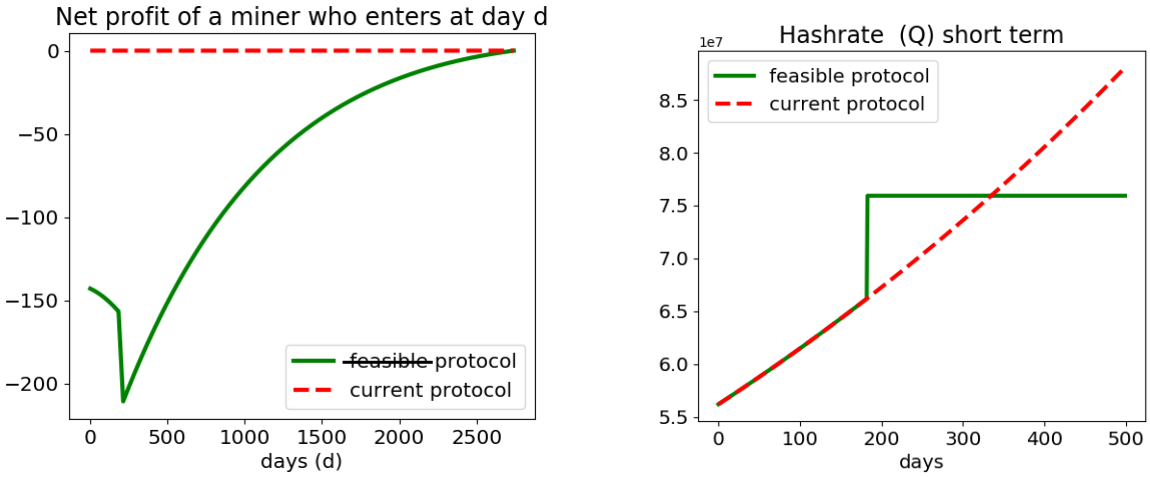
To visualize what the solution looks like, figure 17 compares the optimal feasible protocol with the current one, plotting the protocol design, the hashrate path and the net profit per unit of hashpower miners would make entering the race.

Figure 17: Protocol comparison



The fact that the miner who has the opportunity to invest at time 0 would not make a negative profit investing proves that the incentive-compatibility constraint is satisfied and the fact that the hashrate remains flat at the beginning shows the optimality of the protocol. To see what ill-calibrated protocols look like, figure 18 plots the profit of entering miners for a too greedy protocol ($\overline{B}$ too low) and the hashrate prevailing under a too generous protocol ($\overline{B}$ too high). For the too greedy protocol, we see that the incentive-compatibility constraint is not satisfied (the miner who could enter the market at day $t = 0$ would make losses doing it) and we see that the generous protocol leads to an unnecessary high hashrate.

Figure 18: Too greedy and too generous protocols



Now, to find $\widehat{B}$, I need to know, for a given $\overline{B}$, whether the resulting protocol satisfies the incentive-compatibility constraint or not.

Redefining miners' strategies. So far, I have defined miners' strategies using q_t. Here it is useful to define them differently. Let introduce $\overline{P}_t$, the threshold value for the payoff process at time t, above which investment is triggered. More precisely, if the payoff that would arise at time t without investment, E_t/Q_{t-1}, is bigger than $\overline{P}_t$, then the miner who has the opportunity to invest at time t does it until $E_t/Q_t = \overline{P}_t$. q_t is a function of $\overline{P}_t$. We have:

$$q_t = \max\left(\frac{E_t}{\overline{P}_t} - Q_{t-1}, 0\right) \tag{13}$$

When the miner t chooses q_t, she has $\overline{P}_t$ in mind. So I can alternatively state that miner t's strategy is $\overline{P}_t$. Using this new notation, 12 becomes

$$E_0^{FP}/Q_{-1}^{*CP} \geq \overline{P}_0^{*FP}, \tag{14}$$

The advantage of defining the strategies this way is that the best response of miner t only depends on the strategies of future miners, whereas this is not the case using q_t. As a result, $\left\{\overline{P}_t^*, t \geq t_2\right\}$ is known and is equal to the utopic equilibrium payoff series

since assertion 2.2 ensures that miners invest everyday. Then, the whole equilibrium path $\left\{\overline{P}_t^*,\ t \geq 0\right\}$ can be recovered going backward.

Recovering the equilibrium path $\left\{\overline{P}_t^*,\ t \geq 0\right\}$. Let $t \geq 0$. If miner $t+1$ does not invest at time $t+1$, we have $P_{t+1} = \frac{E_{t+1}}{E_t} P_t$. If she invests, we have $P_{t+1} = \overline{P}_{t+1}^*$. So, P_{t+1} reads

$$P_{t+1} = \min\left(P_t \times \frac{E_{t+1}}{E_t}, \overline{P}_{t+1}^*\right). \tag{15}$$

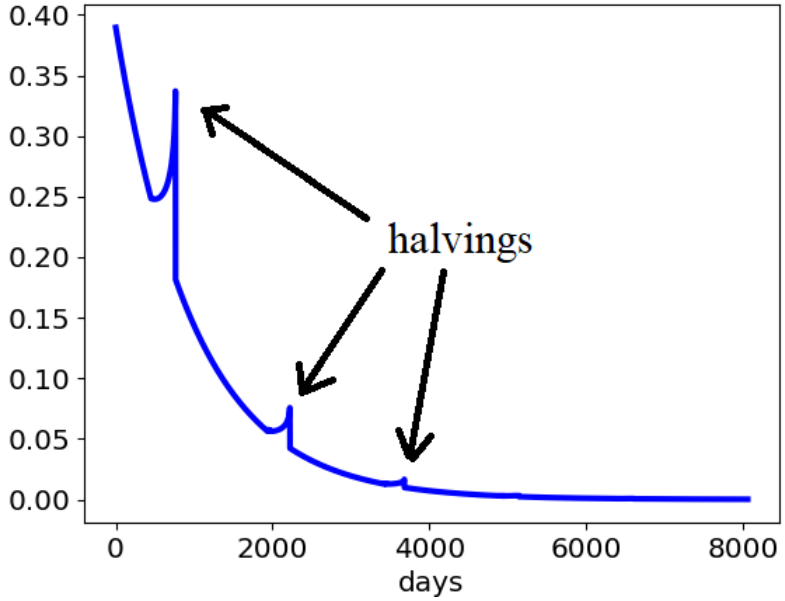
Iterating equation 15 and using the free entry condition, $\overline{P}_t^*$ is the unique solution[41] of the equation (which unknown is P_t)

$$P_t + \sum_{s=t+1}^{+\infty} \frac{\min\left(P_t \times \frac{E_s}{E_t}, \overline{P}_{t+1}^* \times \frac{E_s}{E_{t+1}}, \ldots, \overline{P}_{s-1}^* \times \frac{E_s}{E_{s-1}}, \overline{P}_s^*\right)}{(1+r)^{s-t}} = I_t \tag{16}$$

Using equation 16, $\overline{P}_{t_2-1}^*$ can first be recovered, then $\overline{P}_{t_2-2}^*$, ..., and eventually $\overline{P}_0^*$. To recover Q_{-1}^{*CP}, the same methodology can be applied for the current protocol. The path $\left\{\overline{P}_t^{*CP},\ t \geq 0\right\}$ is displayed on figure 19

[41] the left hand side is clearly an increasing function in P_t.

Figure 19: $\left\{\overline{P}_t^{*CP},\ t \geq 0\right\}$



The curve plotted decreases exponentially (except around halvings) because of technological progress. We see that halvings of the number of newly created coins affect $\overline{P}^{*CP}$ only a couple months ahead, which confirms the findings of PW. As a result, since the date $t = 0$ we consider is far enough from the next halving, we do not need to take halvings into account and Q_{-1}^{*CP} is equal to the equilibrium hashrate prevailing under the hypothetical "flat protocol", where $B_t = 12.5$ forever. This quantity is easily computed using assertion 2.2. The condition in equation 14 can now be checked.

Choice and effects of parameters on the holder's expenditure for the two protocols. The value for $\widehat{B}$ and the magnitude of the efficiency gain switching from the current protocol to the optimal feasible one depend on many parameters. These are the technical progress rate, a, the discount rate, r, the market capitalization at time 0, K_0, the growth rate of the market capitalization, α, the number of available bitcoins at time 0, $\mathcal{B}_0$, the level of fees F and the total costs of one unit of hashpower at time 0: I_0.

The current and the feasible protocols do not differ until date $t = t_1$. Even though the hashrate differs from date $t = 0$ onward, this does not affect the holder's expenditure since it only depends on the protocol design and the exchange rate. So the

comparison starts at date $t = 1$. The holder's expenditure under the current protocol reads

$$Exp^{CP} = K_0 \sum_{t=t_1}^{+\infty} \frac{(1+\alpha)^t B_t^{CP}}{\mathcal{B}_0 + \sum_{s=1}^{t} B_s^{CP}}$$

and the holder's expenditure under the optimal feasible protocol reads

$$Exp^{FP} = K_0 \widehat{B} \sum_{t=t_1}^{t_2-1} \frac{(1+\alpha)^t}{\mathcal{B}_0 + (t-t_1+1)\widehat{B}}.$$

It is clear that Exp^{CP} increases much more than Exp^{FP} with α. Even though α has been very high in the past, for the comparison to be fair, we have to pick a low α. An α significantly higher that the global growth rate is not sustainable in the long term. I adopt a very conservative approach and set $\alpha = 0$, which is the most adverse case (not considering negative values for α) to advocate the protocol change. $\mathcal{B}_0$ is easily observable and I therefore set it at $17,000,000$. All the other parameters may affect Exp^{FP} indirectly through $\widehat{B}$ but they have the same direct effect (or no direct effect at all) on Exp^{CP} and Exp^{FP}. Finally, note that Exp^{FP} is clearly increasing (almost linearly) with $\widehat{B}$. Thus, to assess how the parameters affect the efficiency gain switching to the feasible protocol, it is sufficient to assess their effects on $\widehat{B}$. The lower $\widehat{B}$, the higher the efficiency gain.

The parameter which has by far the strongest effect on $\widehat{B}$ is a. $\widehat{B}$ strongly decreases with a. When a increases, future miners enjoy a much more efficient technology which enables them to invest more. Even under the current protocol, active miners' profits will vanish quickly. So it is not necessary to give them a high compensation for the protocol change. Again, I select an adverse case and set a according to the so-called Moore law, which stipulates that I_t should be divided by two every two years. Note that PW estimate a twice faster technical progress rate. If a is below 10% per year, then the protocol switch becomes pointless because satisfying the incentive-compatibility constraint would make the holder worse-off.

r also has a non-negligible effect on $\widehat{B}$. $\widehat{B}$ decreases with r. When r increases, the future income loss matters less for miners who are thus ready to accept a lower transfer. Following PW, I pick, somewhat arbitrarily, $r = 10\%$ annually.

All the other parameters have negligible effects on $\widehat{B}$. Both the market capitalization and the level of transaction fees are also easily observed but very volatile. I set $K_0 = 200,000,000,000$ \$, which is consistent with a ₿/\$ exchange rate slightly above $10,000$ \$, and $F = 72000$ \$ per day, which is the average fee level. These choices do not matter much since $\widehat{B}$ increases extremely slightly with K_0 and decreases extremely slightly with F. I_0 has no effect at all on $\widehat{B}$. It only affects the hashrate. α also has a negligible negative effect on $\widehat{B}$.

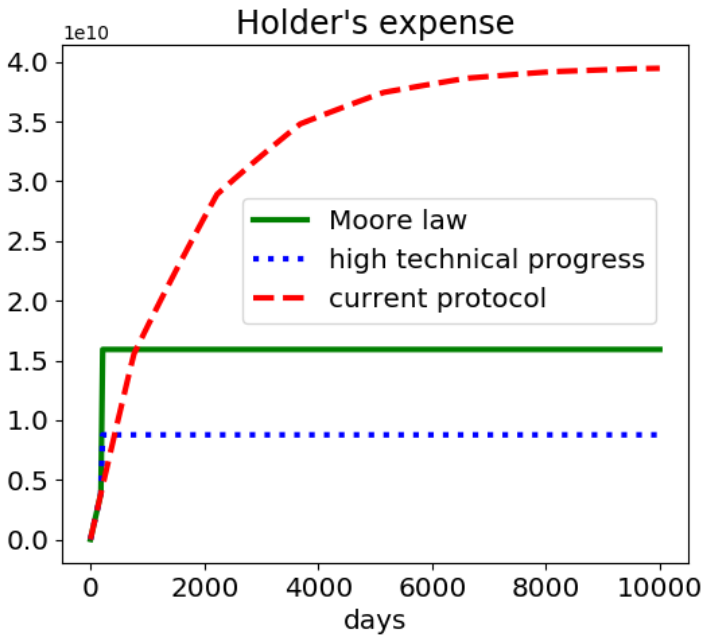
Results Using the values for the parameters previously discussed, I obtain the following optimal feasible protocol.

Optimal feasible protocol.

1. For six months, each new block yields 12.5 new bitcoins.
2. Then for one month, each new block yields 243 new bitcoins.
3. Then, no more new bitcoins are ever created.

Figure 20 compares the holder's total expense between the two protocols, for the Moore law technical progress rate and the rate estimated by PW. As explained in the previous section, we see that under the optimal feasible protocol, the holder is temporarily worse-off but eventually much better-off.

Figure 20: Efficiency gain



To have a quantitative idea of the effects of the different parameters on $\widehat{B}$, table 6 compares the value obtained for $\widehat{B}$ with the default set of parameters, with the $\widehat{B}$ obtained for other sets of parameters, for which one of the parameters is significantly changed (often multiplied by 2). We see that the effect is indeed mostly driven by the technical progress rate.

Table 6: Quantitative effects of parameters

	$\widehat{B}$ (new bitcoins per block)
Default	243
$a \times 2$	99
$a = 10\%$ yearly	649
$r \times 2$	203
$r \approx 0$	299
$K_0 \times 2$	254
$F \times 2$	224
$\alpha = 10\%$ yearly	227

2.5 Conclusion

I have shown that the current Bitcoin protocol is inefficient because holders pay a tax which is used to buy hardware and electricity without any gain for the security of transactions. Switching to the first best protocol is unfortunately impossible in practice because it would harm miners, who have the power to prevent any protocol change. I have then pined down numerically the second best protocol, which is basically the first best protocol with a transfer to miners just high enough to satisfy the incentive-compatibility constraint. Finally, I have shown that the magnitude of the efficiency gain offered by the second best protocol depends crucially on the technical progress rate as regard mining hardware. Bitcoin holders will eventually benefit implementing the protocol change provided the yearly technological progress rate is above 10%.

3 Identification and estimation of the average marginal effect in a panel data fixed effects logit model

Joint work with Laurent Davezies and Xavier D'Haultfœuille.

Abstract

This article focuses on the average marginal effect in a fixed effects logit model with panel data. We show that this quantity is partially identified and provide a nice and easy to use characterization for the sharp bounds of the identification region. We then explain how to estimate those bounds, using our characterization. Finally, we perform Monte-Carlo simulations to assess the performance of our estimators.

3.1 Introduction

So far no method is truly satisfying when drawing inference on a binary variable using panel data. Panel data can help mitigate the endogeneity problem since they enable the econometrician to decompose the error term into an individual effect, fixed in time, and an idiosyncratic shock. For many models, the data can be transformed so as to get rid of the individual effect. Thus, only the indiosyncratic shock needs not be correlated with the regressors and no assumption must be made on the individual effect and its link with the regressors. Note that the individual effects cannot be treated as parameters and estimated due to the, so-called, incidental parameter problem. See Lancaster (2000) for a survey on this issue. In this case, when the number of individuals increases, the number of parameters to estimate increases at the same speed. This phenomenon prevents the econometrician from correctly estimating not only the individual effects, but also the parameter of interest. If one wants to avoid making assumptions on the individual effects, one must get rid of them.

The simplest way to proceed is to assume a linear model. Simply differentiating the data makes the individual effects disappear. However linear models with binary dependent variables (called linear probability models) have a well-known drawback: they imply that the marginal effect of a regressor on the explained variable is the same for all individuals, no matter what their unobserved individual components are. This

is a serious restriction. On top of that, should the model be used for prediction, linear probability models do not restrict the probability that the explained variable equal one to belong to the set $[0,1]$. Fortunately, one can also assume a, so-called, fixed effect logit model. Rasch (1961) and Chamberlain (1980) have found an exhaustive statistic for the individual effect (called the fixed effect) in this model. Proceeding conditionally to this statistic, the data do not depend on the fixed effect anymore and this enables the econometrician to consistently estimate β_0, the parameter of interest.

Yet, in a logit model, each individual has its own marginal effect, which depends on his fixed effect. Even though this is what we wanted, it complicates the estimation. Different quantities can sum up the effect for the whole population. In this article, we focus on the average marginal effect. Knowing β_0 is, however, not enough to recover the average marginal effect because it depends on the unknown distribution of the fixed effects. The model still yields some information on that distribution but not enough to point-identify the average marginal effect, which remains partially identified. Chernozhukov et al. (2013) give a first and natural characterization of the sharp bounds for the average marginal effect. The characterization they give induces an estimation strategy which involves a maximization over an infinite-dimensional set. Such a problem is very difficult to solve both in practice and in theory.

In this article, we provide another characterization of the sharp bounds, relying on the theory of Chebyshev systems, using many results from Krein and Nudel'man (1977). Our characterization is nicer since it involves only some costless operations, namely solving a linear system and finding the roots of a polynomial. We proceed almost as in D'Haultfœuille and Rathelot (2011), who resort to the same methodology to study segregation indexes.

This article builds on the partial identification literature. This notion was democratized by Manski (2003). It generalizes point-identification and enables the econometrician to still extract information on a parameter from the data while relaxing some restrictive assumptions needed in classical models. Many articles in this literature yield sharp bounds in very general frameworks. But often, this generality comes with great difficulties to apply the method in practice. In Chesher et al. (2013) and Galichon and Henry (2011), for instance, the econometrician obtains an infinite number of restrictions. We rather do the opposite: the framework we consider is more specific and enables us to provide a very easy-to-use characterization of the sharp identification bounds. Thus, the philosophy of our article is closer to Davezies and D'Haultfœuille

(2016), Bontemps et al. (2012) and Kaido and Santos (2014) who use convexity restrictions to derive simple results.

The rest of the article is organized as follows: section 2 presents the identification results, section 3 deals with the estimation procedure, section 4 shows some Monte-Carlo simulations and section 5 concludes.

3.2 Identification results

3.2.1 The parameters of interest

We suppose we observe panel data of the form $(Y_1, X_1, ..., Y_T, X_T)$, with $Y_t \in \{0,1\}$ and $X_t = (X_{1t}, ..., X_{pt}) \in \mathbb{R}^p$ for $t \in \{1, ..., T\}$. Note that in this section, we do not index units by i for the sake of lighter notations. We focus on the so-called fixed effect logit by considering the following assumption.

Assumption 9. *We have*

$$Y_t = \mathbb{1}\{X_t\beta_0 + \alpha + \varepsilon_t \geq 0\},$$

where α, ε_1,...,ε_T *are real random variables. Moreover,* $(\varepsilon_1, ..., \varepsilon_T)$ *are i.i.d., follow a logistic distribution and are independent of* $(\alpha, X_1, ..., X_T)$.

Importantly, this assumption allows for arbitrary dependence between the individual effect α and the covariates $X = (X_1, ..., X_T)$. We are interested here in the identification of aggregate marginal effects. Without loss of generality, we focus on marginal effects on the last period. Assumption 9 implies that $\Pr(Y_T = 1|X_T, \alpha) = \Lambda(X_T\beta_0 + \alpha)$, with $\Lambda(x) = 1/(1 + \exp(-x))$, the c.d.f of the logistic distribution. Thus, the marginal effect of X_{kT} on Y_T is $\beta_{0k}\Lambda'(X_T\beta + \alpha)$, with Λ' the derivative of Λ. Rather than such individual effects, we focus here on an aggregate measure, namely

$$\Delta_k = \beta_{0k}\mathbb{E}\left[\Lambda'(X_T\beta_0 + \alpha)\right]$$

In the following, we show that this parameter is partially identified, and provide a useful characterization of the corresponding bounds. This characterization will, in turn, allow us to produce a fairly simple estimator. Before, let us recall that β_0 is identified in this model, as the maximizer of the (expected) conditional log-likelihood. See Rasch (1961) and Chamberlain (1980). For completeness, we first state this result below.

Let $S = \sum_{t=1}^{T} Y_t$, let $C_k(x;\beta) = \sum_{\substack{(d_1,\ldots,d_T)\in\{0,1\}^T: \\ \sum_{t=1}^{T} d_t = k}} \exp\left(\sum_{t=1}^{T} d_t x_t \beta\right)$ and let $\ell_C(y_1,...,y_T|k,x;\beta) = \sum_{t=1}^{T} y_t x_t \beta - \ln\left(C_k(x;\beta)\right)$.

Proposition 3.1. *Suppose that Assumption 9 holds. Then*

$$\beta_0 = \arg\max_{\beta} \mathbb{E}\left[\ell_C\left(Y_1,...,Y_T|S,X;\beta\right)\right].$$

Proof. See Chamberlain (2010). □

Because β is identified, computing bounds on Δ reduces to the identification of features of the distribution of α conditional on X. To see this, note that

$$\Delta_k = \beta_{0k}\mathbb{E}\left[\mathbb{E}\left[\Lambda'(X_T\beta_0 + \alpha)|X\right]\right]$$

This equality implies that it is sufficient to find bounds on the prescribed moment of α given X. Specifically, one must bound, $\Delta(X) = \mathbb{E}\left[\Lambda'(X_T\beta_0 + \alpha)|X\right]$.

3.2.2 Sharp bounds

We now show that $\Delta(X)$ is only partially identified in general. Conditionally on X, all the information contained in the data is

$p_X(y_1,\ldots,y_T) \equiv \Pr\left(Y_1 = y_1,\ldots,Y_T = y_T|X\right)$. Integrating out the conditional distribution of α, we see that those probabilities read

$$p_X(y_1,\ldots,y_T) = \int \prod_{t=1}^{T} \Lambda(X_t'\beta_0 + \alpha)^{y_t}\left(1 - \Lambda(X_t'\beta_0 + \alpha)\right)^{1-y_t} dF(\alpha|X), \quad (y_1,\ldots,y_T) \in \{0,1\}^T \tag{17}$$

So the model provides us with some constraints on the conditional distribution of α because the left hand side of this equality is identified in the data. We know some moments of this distribution but not the moment corresponding to $\Delta(X)$, in general. In some cases we will pin down below, $\Delta(X)$ is a linear combination of the identified moments. Yet, this is not always the case. Thus, in general, $\Delta(X)$ is only partially identified. Finding the sharp lower bound (resp. upper bound) of its identification region amounts to minimizing (resp. maximizing) $\int \Lambda'\left(X_T'\beta_0 + \alpha\right) dF(\alpha)$ over the set of all

distributions F that satisfy the constraints (17). Those constraints encompass all the available information so the bounds will indeed be sharp.

We now provide a first characterize of the bounds. The idea is to obtain constraints on raw moments[42] of some random variable instead of the complicated moments we now have. For that purpose, let us introduce some additional notation. First, let $A = (a_{i,j})_{0\leq i,j\leq T} = \binom{T-i}{j-i} I(i \leq j)$ and

$$(c_0(x), \ldots, c_T(x))' = A \left(\frac{\Pr(S=0|X=x)}{C_0(x;\beta_0)}, \ldots, \frac{\Pr(S=T|X=x)}{C_T(x;\beta_0)} \right)'.$$

Second, for $k = 0, \ldots, T$, let

$$m(x) = (m_1(x), \ldots, m_T(x)) = \left(\frac{c_1(x)}{c_0(x)}, ..., \frac{c_T(x)}{c_0(x)} \right). \tag{18}$$

Note that because $m(x)$ only involves β_0 and the conditional distribution of S, it is identified from the data. Third, let us define

$$\Omega(u,x) = \frac{u(1-u)\prod_{t=1}^{T-1}(u(\exp(x_t\beta)-1)+1)}{u(\exp(x_T\beta)-1)+1}.$$

Finally, let $\mathcal{D}$ denote the set of cumulative distribution functions on $[0,1]$ and for any $m = (m_1, ..., m_T)$, let

$$\mathcal{F}(m) = \left\{ F \in \mathcal{D} : \int u^t dF(u) = m_t,\ t = 1, \ldots, T \right\}.$$

Proposition 3.2. *The sharp identification region of* $\Delta(x)$ *is* $\left[\underline{\Delta}(x), \overline{\Delta}(x)\right]$*, with*

$$\underline{\Delta}(x) = \exp(x_T\beta)c_0(x) \min_{F\in\mathcal{F}(m(x))} \int \Omega(u,x)dF(u),$$
$$\overline{\Delta}(x) = \exp(x_T\beta)c_0(x) \max_{F\in\mathcal{F}(m(x))} \int \Omega(u,x)dF(u).$$

Moreover, $\underline{\Delta}(x) = \overline{\Delta}(x)$ *whenever* $x_t = x_T$ *for some* $t < T$.

Proof. Using a simple change of variable, the objective function becomes:

$\Delta(X) = \exp(X_T\beta_0)\int_0^1 \frac{u(1-u)}{\left(u(\exp(X_T'\beta_0)-1)+1\right)^2} dF(u|X)$, where $F(.|X)$ is the conditional distribution of $U = \frac{\exp(\alpha)}{1+\exp(\alpha)}$ given X.

[42] for a random variable U, raw moments are $E[U]$, $\mathbb{E}[U^2]$,.... We do not call them canonical moments because this terminology has a very specific meaning in the theory we use.

As for the constraints on the conditional distribution of α, we actually do not need the 2^T of them. Since S is a sufficient statistic for α, all the relevant information for our application is encompassed in the $T+1$ conditional probabilities $\Pr(S=0|X), \ldots, \Pr(S=T|X)$.[43]

For all $k=0,\ldots,T$, we have

$$\Pr(S=k|X) = C_k(X,\beta_0) \int \frac{\exp(k\alpha)}{\prod_{t=1}^{T}(1+\exp(X_t'\beta_0+\alpha))} dF(\alpha|X).$$

The same change of variables as before yields

$$\Pr(S=k|X) = C_k(X,\beta_0) \int_0^1 \frac{u^k(1-u)^{T-k}}{\prod_{t=1}^{T}(u(\exp(X_t'\beta_0)-1)+1)} dF(u|X). \tag{19}$$

Using the notation we have defined, we can rewrite the set of constraints as:

$$\text{for all } k=0,\ldots,T,\ \ c_k(X) = \int_0^1 \frac{u^k}{\prod_{t=1}^{T}(u(\exp(X_t'\beta_0)-1)+1)} dF(u|X).$$

In order to get rid of the denominator so that our constraints are raw moments, we introduce a new conditional measure $G(u|X)$ such that $\frac{dF(u|X)}{dG(u|X)} = c_0(X)\prod_{t=1}^{T}(u(\exp(X_t'\beta_0)-1)+1)$. The constraints now become

$$\begin{cases} \int_0^1 dG(u|X) &= 1 \ \text{(G is a probability measure on [0,1])} \\ \int_0^1 u^k dG(u|X) &= m_k(X),\ \ k=1,\ldots,T \end{cases} \tag{20}$$

And $\Delta(X)$ becomes

$$\Delta(X) = \exp(X_T\beta_0)c_0(X) \int_0^1 \Omega(u,X)\, dG(u|X)$$

We have found the expression of proposition 3.2. Before all the transformations we did in this proof we had sharp bounds. Nowhere did we lose information in the process. So the bounds of proposition 3.2 are still sharp.

Finally, if $x_t = x_T$ for some t, then one of the terms in the product in the numerator of $\Omega(u,x)$ will cancel with the denominator, leaving a polynomial of degree at most T. Since the identified moments (the constraints) define a basis of the vector space of polynomials with degree at most T, $\Delta(X)$ can be expressed as a linear combination of the identified moments. □

[43] Note that this is the property we use to identify β_0.

This proposition first shows that finding the sharp bounds on $\Delta(x)$, and thus on Δ, amounts to bounding a prescribed moment of a distribution on $[0,1]$, given the knowledge of its first T raw moments. It also shows that the conditional distribution of S, as opposed to the one of $(Y_1,...,Y_T)$, is sufficient to obtain sharp bounds on $\Delta(x)$. This feature is convenient for estimation, because for T large, the $T+1$ probabilities of the distribution of S will be estimated more accurately than the 2^T probabilities of the distribution of $(Y_1,...,Y_T)$. Finally, the proposition shows that the bounds are actually equal for "stayers", that is to say units satisfying $x_t = x_T$ for some $t < T$. This result was expected and is not specific to the logit model we consider here. It is indeed a particular case of Hoderlein and White (2012).

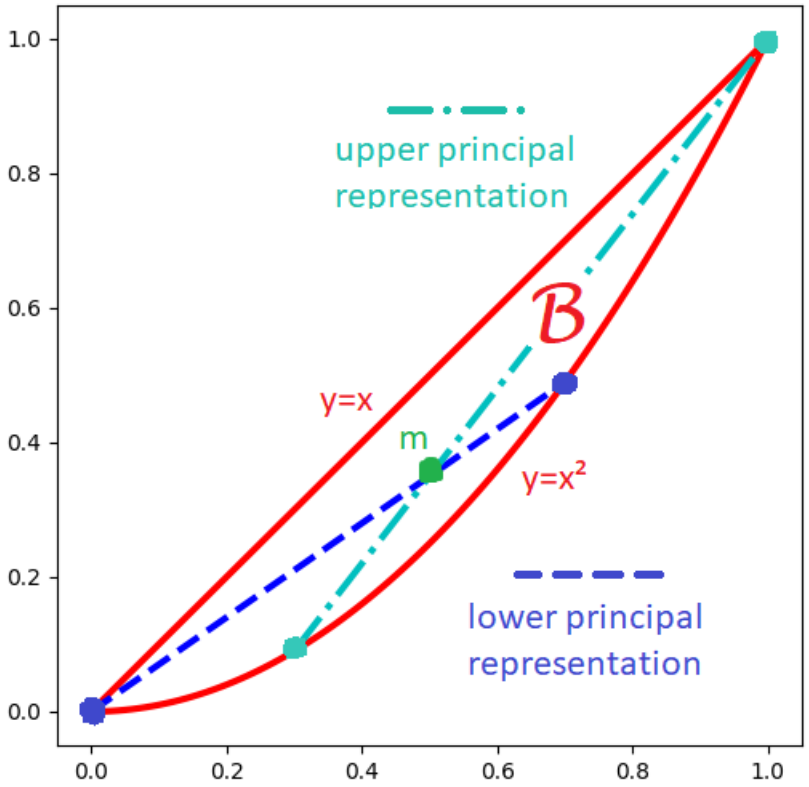
Now, the problem of finding bounds of a moment given other moments can be difficult. Formulated as in Proposition 3.2, it involves an infinite dimensional constrained optimization. Several simplifications are however possible. First, an application of Caratheodory's theorem shows that one can actually restrict the set of potential distributions to discrete distributions with at most $T+1$ support points. This turns the problem into a nonlinear constrained optimization problem of dimension $2T+1$. Such an optimization problem can still be difficult to solve, yet. Fortunately, further simplifications are possible when the function Ω involved in the moment $\int \Omega(u,x)dF(u)$ under consideration satisfies certain conditions that are detailed below. For us those conditions are satisfied. As a result, special and discrete distributions, called the lower and upper principal representations, rationalize the moment constraints and yield the sharp bounds on the prescribed moment. Moreover, those two distributions are very parsimonious and can be computed very easily.

So far, the bounds are defined as the solutions of two optimization problems, for which non convex optimization is to be performed over infinite-dimensional sets. Solving such problems is usually very difficult. This is where lies the main added value of this paper: we provide a very simple and computationally almost costless way to do so, which gives us a much nicer characterization of the two bounds.

3.2.3 A nicer characterization for the bounds

Before stating the main result, we illustrate the intuition in the two-period case. Let $\mathcal{B} \equiv \{m \in [0,1]^2 \text{ such that } \mathcal{F}(m) \neq \varnothing\}$. $\mathcal{B}$ is the set of moment constraints that can be rationalized for some conditional distribution for the transformed fixed effects $U \equiv \Lambda(\alpha)$. In the two-period case it is easy to see that $\mathcal{B}$ is the area between the segment $\{y = x \text{ for } x, y \in [0,1]\}$ and the curve $\{y = x^2 \text{ for } x, y \in [0,1]\}$. $\mathcal{B}$ is represented on figure 21.

Figure 21: Graphical example with two periods



If the model is true, then the vector $m(x) = (m_1(x), m_2(x))$ corresponds to the two first moments of a random variable supported by $[0,1]$. So we must have $m(x) \in \mathcal{B}$. If $m(x)$ is on the lower border of $\mathcal{B}$ (say, for example, $m_1(x) = p \in [0,1]$ and $m_2(x) = p^2$, then the only distribution F for U that can rationalize the constraints is the Dirac distribution centered at p. Indeed it implies that $\int u dF(u) = p$ and $\int u^2 dF(u) = p^2$ so $\mathbb{V}(U) = 0$. If $m(x)$ is on the upper border of $\mathcal{B}$, say, $m_1(x) = m_2(x) = p$, then the only

distribution $F()$ for U which can rationalize the constraints is the Bernoulli distribution with parameter p. In those two particular cases, the distribution of U is point identified and so is $\Delta(x)$. If $m(x)$ is in the interior of $\mathcal{B}$, then the set $\mathcal{F}(m(x))$ contains an infinite number of elements and the conditional distribution of U is partially identified. Yet, the distributions we are looking for (the one which minimizes $\int \Omega(u,x)dF(u)$ and the one which maximizes that same quantity) can be obtained very easily!

Given the convexity of $\mathcal{B}$,[44] as shown on figure 21, if $m(x)$ is in the interior of $\mathcal{B}$, it can be obtained as a linear combination of two points on the lower border of $\mathcal{B}$. Such a linear combination actually defines a distribution with two support points since a point on the lower border corresponds to a Dirac distribution. We focus on two specific combinations: the one where one of the two points on the lower border is $(0,0)$ and the one where one of the two points is $(1,1)$. Those two distributions are called the lower and upper principal representation respectively and are plotted on figure 21. It turns out that one of those two distributions is the one which minimizes $\int \Omega(u,x)dF(u)$ and the other one maximizes that same quantity. Not only are those two distributions easy to obtain, they also make the computation of $\int \Omega(u,x)dF(u)$ straightforward since they turn the integral into a sum of two terms! We now introduce the prerequisite needed to state the main result.

Definition 3.3 (Chebyshev systems)**.** *A set of $K+1$ functions is called a Chebyshev system on an interval $[a;b]$ if any non-trivial linear combination of these $K+1$ functions vanishes at K different points of $[a;b]$ at most.*

Lemma 3.1. *If for all $t < T$, $\exp(x_t\beta_0) \neq \exp(x_T\beta_0)$, then the set of $T+2$ functions (which variable is "u") $\left\{1, u, u^2, \ldots, u^T, \frac{u(1-u)\prod_{t=1}^{T-1}(u(\exp(x_t'\beta_0)-1)+1)}{u(\exp(x_T'\beta_0)-1)+1}\right\}$ is a Chebyshev system on $[0,1]$.*

Remark 3.1. *We have already seen that if $\exp(x_t\beta_0) = \exp(x_T\beta_0)$ for some $t < T$, then the marginal effect conditional to x is identified. This is the simplest but not the most interesting case.*

Proof. Multiplying all the functions of a Chebyshev system by the same strictly positive function preserves the Chebyshev property.[45] Let multiply all our functions by

[44] It is obvious that $\mathcal{B}$ is convex. Indeed, if $m^1 \in \mathcal{B}$, then there exist a distribution F_1 with support $[0,1]$ such that $\int udF_1(u) = m_1^1$ and $\int u^2 dF_1(u) = m_2^1$. Idem for some m^2 and F_2. Then, the moment vector $(m^1+m^2)/2$ belongs also to $\mathcal{B}$ since it is rationalized by the distribution $(F_1+F_2)/2$.

[45] Indeed, let g be a strictly positive function. We have: $\sum_{k=1}^K \lambda_k f_k(x) = 0 \Leftrightarrow \sum_{k=1}^K \lambda_k g(x) f_k(x) = 0$

$u(\exp(x_T'\beta_0)-1)+1>0$ for $u\in[0,1]$. Proving the lemma is thus equivalent to showing that the set of functions
$\Big\{u(\exp(x_T\beta_0)-1)+1,\dots,u^T(u(\exp(x_T\beta_0)-1)+1),$

$u(1-u)\prod_{t=1}^{T-1}(u(\exp(x_t\beta_0)-1)+1)\Big\}$ is a Chebyshev system. The $T+1$ functions $u^k(u(\exp(x_T'\beta_0)-1)+1)$, $k=0,\dots,T$ are clearly linearly independent. Any linear combination of these function can be written as $(u(\exp(x_T\beta_0)-1)+1)\sum_{k=0}^{T}a_ku^k$. The last function,
$u(1-u)\prod_{t=1}^{T-1}(u(\exp(x_T\beta_0)-1)+1)$, cannot be written like this since for all $t<T$, $\exp(x_t\beta_0)\neq\exp(x_T\beta_0)$. Thus, we have $T+2$ linearly independent polynomials of degree at most $T+1$. This is a base of the set of polynomials of degree at most $T+1$. No non trivial linear combination of those functions can vanish more than $T+1$ times. □

We need to (re-)define a few more notions in order to state the main result in a very general case. First, let $f_0,\dots,f_T$ be some functions defined on $[0,1]$ and let

$$\tilde{F}(m)=\left\{F\in\mathcal{D}\text{ such that }\int f_k(u)dF(u)=m_k,\ \ k=0,\dots,T\right\}$$

In our case, $f_k(u)=u^k$ and $\tilde{F}=F$. Let also

$$\tilde{\mathcal{B}}=\left\{m\in[0,1]^T\text{ such that }\tilde{F}(m)\neq\varnothing\right\}$$

$\tilde{\mathcal{B}}$ is the set of values for which the moment constraints can be rationalized by some distribution on $[0,1]$. Second, for a point m in $\tilde{\mathcal{B}}$, any distribution in $\tilde{\mathcal{F}}(m)$ is called a *representation* of the moment vector m. We know that if m is in the interior of $\tilde{\mathcal{B}}$, there exists an infinity of representations. Most of them are not discrete. Yet, in the two-period case we saw that the solution to our problem was obtained for a discrete distribution. This result remains true in the general case so we only consider discrete representations. Last, the number of support points of a discrete representation is called the *index* of the representation, with the exception that 0 and 1 are counted only $1/2$ if they are support points.[46]

Let $\tilde{\Omega}$ be some function defined on $[0,1]$. We focus on the following problem: bound the quantity $\int\tilde{\Omega}(u)dF(u)$, over the set $\tilde{\mathcal{F}}(m)$.

[46]for example, the index of a representation where 0.3 and 0.7 are support points is 2, the index of a representation where 0 and 0.6 are support points is $3/2$ and the index of a representation where 0 and 1 are support points is 1.

We can now state the main result which tells us how to find the sharp bounds for the marginal effect.

Theorem 3.4 (Main result)**.**
Assume that the set of functions $(f_0, \dots, f_T)$ *is a Chebyschev system. Then,*
(1) A point m *is on the border of* $\tilde{\mathcal{B}}$ *if and only if it admits a representation of index at most* $\frac{T}{2}$*. In this case, this representation is unique.*
(2) A point m *is in the interior of* $\tilde{\mathcal{B}}$ *if and only if it admits no representation of index strictly smaller than* $\frac{T+1}{2}$ *and exactly two representations of index* $\frac{T+1}{2}$*. Those two representation are called the principal representations of the point* m*.*[47]
Now assume, on top, that m *is in the interior of* $\tilde{\mathcal{B}}$*.*
(2') If T *is even, one of the representation has the support point* 0 *(it is called the lower principal representation) and the other representation has the support point* 1 *(it is called the upper principal representation). If* T *is odd, then one of the representations has neither the support point* 0*, neither the support point* 1 *(it is called the lower principal representation) and the other representation has both the support points* 0 *and* 1 *(it is called the upper principal representation).*
(2") Let F_- *and* F_+ *denote the lower and upper principal representations of* m *respectively and let assume that the set of functions* $(f_0, \dots, f_T, f_{T+1} \equiv \tilde{\Omega})$ *is a Chebyshev system. If for all* $0 < u_0 < \cdots < u_{T+1} < 1$*, the determinant of the matrix* $(f_i(u_j))_{0 \leq i,j \leq T+1}$ *is strictly positive, then, we have* $\int \tilde{\Omega}(u) dF_-(u) = \min_{F \in \tilde{\mathcal{F}}(m)} \int \tilde{\Omega}(u) dF(u)$ *and* $\int \tilde{\Omega}(u) dF_+(u) = \max_{F \in \tilde{\mathcal{F}}(m)} \int \tilde{\Omega}(u) dF(u)$*. If for all* $0 < u_0 < \cdots < u_{T+1} < 1$*, the determinant of the matrix* $(f_i(u_j))_{0 \leq i,j \leq T+1}$ *is strictly negative, then, we have* $\int \tilde{\Omega}(u) dF_+(u) = \min_{F \in \tilde{\mathcal{F}}(m)} \int \tilde{\Omega}(u) dF(u)$ *and* $\int \tilde{\Omega}(u) dF_-(u) = \max_{F \in \tilde{\mathcal{F}}(m)} \int \tilde{\Omega}(u) dF(u)$*.*

Proof. See Krein and Nudel'man (1977), chapter 4, theorem 1.1, page 109 □

Remark 3.2. *The fact that for all* $0 < u_0 < \cdots < u_{T+1} < 1$*, the sign of the determinant of the matrix* $(f_i(u_j))_{0 \leq i,j \leq T+1}$ *is constant is ensured by* $(f_0, \dots, f_T, f_{T+1} \equiv \tilde{\Omega})$ *being a Chebyshev system.*

Note that the two principal representations do not depend on $\tilde{\Omega}$. Therefore the same principal representations can be used to find bounds for the quantity $\int \tilde{\Omega}(u) dF(u)$, with a different $\tilde{\Omega}$ as long as those functions define a Chebyshev system. This is a very powerful result! This theorem applies in our case. Indeed, it is obvious that the set of

[47] Of course, it admits many representations of greater indexes, which we do not care about here.

functions $(1, u, u^2, \ldots, u^T)$ is a Chebyshev system and we have checked that this property is conserved when the function $\Omega(., x)$ is added to the set. We now explain how to find the two principal representations.

3.2.4 Computation of the bounds

For the sake of simpler notations, we omit the dependency on x in "$m(x)$" for this part. If m belongs to the border of $\mathcal{B} \equiv \left\{m \in [0,1]^T \text{ such that } \mathcal{F}(m) \neq \varnothing\right\}$, the marginal effect is identified since it is a linear combination of the identified moment constraints. Computing it is straightforward. From now on we assume that m is in the interior of $\mathcal{B}$. Once the two principal representations are known, computing the bounds is easy because the integral $\int \Omega(u, x) dF(u)$ becomes a sum of just a few terms. The only challenge is to get the two principal representations. Although at this point it is not obvious how to get them, once known, the procedure is very simple and computationally almost costless for it involves only inverting a linear system and finding the roots of a polynomial. Indeed, we start with giving ourselves a polynomial, which roots are the support points. The restrictions on those, given by theorem 3.4, enable us to obtain a linear system, which unknowns are the coefficients of the polynomial. Once we have the support points, we can easily write another linear system, which unknowns are the weights associated to the support points. We detail the procedure for T even and for the upper principal representation only. The three other cases are very similar and the procedure can be readily adapted.

We know that there are $\frac{T}{2}+1$ support points, including 1. So, let $P = (1-X)\sum_{k=0}^{T/2} e_k X^k$ be a polynomial of degree $\frac{T}{2}+1$ whose roots are the support points. The $T+1$ moment conditions give us equations to recover the coefficients e_k, $k = 0, \ldots, T/2$.

Let F_+ denote the upper principal representation. Let $(p_1, \ldots, p_{T/2+1})$ denote its support points and let $(w_1, \ldots, w_{T/2+1})$ denote the associated weights. Since F_+ is a representation of m, it satisfies the moment conditions. So we have

$$\mathbb{E}^{F_+}\left[U^t\right] = m_t, \;\; t = 0, \ldots, T, \text{ with } m_0 \equiv 1. \tag{21}$$

Using equation 21, we can show that

$$\text{for } l = 0, \dots, \frac{T}{2} - 1, \quad \sum_{k=0}^{T/2} e_k \left(m_{l+k} - m_{l+k+1} \right) = 0. \tag{22}$$

Indeed,

For $l = 0, \dots, \frac{T}{2} - 1$, we have:

$$\begin{aligned}
\sum_{k=0}^{T/2} e_k \left(m_{l+k} - m_{l+k+1} \right) &= \sum_{k=0}^{T/2} e_k \mathbb{E}^{F_+} \left[U^{l+k}(1-U) \right] \\
&= \mathbb{E}^{F_+} \left[U^l (1-U) \sum_{k=0}^{T/2} e_k U^k \right] \\
&= \mathbb{E}^{F_+} \left[U^l P(U) \right] \\
&= \sum_{j=1}^{T/2+1} w_j p_j^l P(p_j) \\
&= \sum_{j=1}^{T/2+1} w_j p_j^l \times 0 \\
&= 0.
\end{aligned}$$

Multiplying P by a non-zero constant does not change its roots. We can thus set $e_{T/2} = 1$. The set of equation 22 becomes

$$m_{l+\frac{T}{2}+1} - m_{l+\frac{T}{2}} = \sum_{k=0}^{T/2-1} e_k \left(m_{k+l} - m_{k+l+1} \right) \ , l = 0, \dots, \frac{T}{2} - 1. \tag{23}$$

The set of equation 23 defines a system, which unknowns are the coefficients of P:

$$\begin{pmatrix}
1 - m_1 & m_1 - m_2 & \cdots & m_{\frac{T}{2}-1} - m_{\frac{T}{2}} \\
m_1 - m_2 & m_2 - m_3 & \cdots & m_{\frac{T}{2}} - m_{\frac{T}{2}+1} \\
\vdots & \vdots & & \vdots \\
m_{\frac{T}{2}-1} - m_{\frac{T}{2}} & m_{\frac{T}{2}} - m_{\frac{T}{2}+1} & \cdots & m_{T-2} - m_{T-1}
\end{pmatrix}
\begin{pmatrix}
e_0 \\ e_1 \\ \vdots \\ e_{\frac{T}{2}-1}
\end{pmatrix}
=
\begin{pmatrix}
m_{\frac{T}{2}+1} - m_{\frac{T}{2}} \\
m_{\frac{T}{2}+2} - m_{\frac{T}{2}+1} \\
\vdots \\
m_T - m_{T-1}
\end{pmatrix}$$

Simply solving the system yields the support points. As for the vector of weights, it is also the solution of a linear system.

Using only the $T/2$ first moments, we obtain:

$$m_k = \mathbb{E}^{\bar{F}}\left[U^k\right] = \sum_{j=1}^{T/2+1} w_j p_j^k = \sum_{j=1}^{T/2} w_j p_j^k + \left(1 - \sum_{j=1}^{T/2} w_j\right) p_{T/2+1}^k \tag{24}$$

Rearranging equation 24, we get:

$$m_k - p_{T/2+1}^k = \sum_{j=1}^{T/2} w_j \left(p_j^k - p_{T/2+1}^k\right) \quad k = 1, \ldots, T/2 \tag{25}$$

So the weights are obtained solving the system:

$$\begin{pmatrix} p_1 - p_{T/2+1} & \cdots & p_{T/2} - p_{T/2+1} \\ \vdots & & \vdots \\ p_1^{T/2} - p_{T/2+1}^{T/2} & \cdots & p_{T/2}^{T/2} - p_{T/2+1}^{T/2} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_{T/2} \end{pmatrix} = \begin{pmatrix} m_1 - p_{T/2+1} \\ \vdots \\ m_{T/2} - p_{T/2+1}^{T/2} \end{pmatrix}$$

We have shown that in general the average marginal effect is partially identified, we have derived sharp bounds for the identification interval as well as a very nice characterize of the bounds. Finally, we have shown how those bounds can be computed. The last point - computation of the bounds - is still part of the "identification" section. There is no data so far. Indeed some problems can (do) arise when data are used.

3.3 Estimation

3.3.1 The use of an index

The "identification" part suggests us to estimate $\Delta(x)$ for all x in the support of X and then integrate the conditional effects with respect to the empirical distribution of X. Processing this way might however not be a brilliant idea since the first step of the estimation procedure is to estimate $\Pr(\sum_{t=1}^{T} Y_t = k|X = x)$ nonparametrically. If there are a couple covariates, such an estimator may not behave very well at all due to the curse of dimensionality because the dimension of X is $k \times T$. Since in the expression of the parameter of interest, $\Delta(x)$, and the constraints, $\Pr\left(\sum_{k=0}^{T} Y_t = k|X = x\right)$, there is always $V_t \equiv \exp(X_t\beta_0)$ and never X alone, it is possible to do the whole machinery conditional to $V = (V_1, \ldots, V_T)$ instead of X. Now, the dimension of the regressor in the nonparametric regression is T instead of $k \times T$. Of course, proceeding that way we obtain outer bounds because we do not use all the information available in the

data. Yet at the end of the day it is not the size of the identification region that matters but the size of the confidence region. In most cases, using the index V will probably strongly decrease the variance of the estimators of the bounds while barely increase the size of the identification region. So, after estimating β, our first step is to estimate $\Pr(S = k|V = v)$, $k = 0, \dots, T$. Then, we get $\hat{m}(v)$ substituting $\hat{\Pr}(S = k|V = v)$ to $\Pr(S = k|V = v)$, $k = 0, \dots, T$ in equation 18.

We know that if the model is true, $m(v)$ must be in $\mathcal{B}$. Yet, nothing ensures that $\hat{m}(v)$ will belong to $\mathcal{B}$! This issue gets more severe when T increases because the volume of the set $\mathcal{B}$ shrinks exponentially fast with T (see Gamboa and Lozada-Chang (2004)). If $\hat{m}(v) \in \mathcal{B}$, we can proceed exactly as in the "identification" section. But in the general case, this procedure is not available any more when $\hat{m}(v) \notin \mathcal{B}$. To overcome this issue, there are at least two solutions. The first idea essentially boils down to projecting $\hat{m}(v)$ on $\mathcal{B}$. The second idea, available only in the two-period case, is to always follow the procedure we follow when $m(v) \in \mathcal{B}$.

3.3.2 A first idea

When $\hat{m}(v) \notin \mathcal{B}$ it seems natural to project $\hat{m}(v)$ on $\mathcal{B}$. Projecting with respect to the euclidean distance is very difficult because it involves finding out the exact geometry of $\mathcal{B}$ for any number of periods T. Instead, we can resort to a procedure close to maximum likelihood. We now estimate $m(v)$ under the constraint that $\hat{m}(v) \in \mathcal{B}$. Let $q_k(v) = \Pr(S = k|V = v)$. Conditional to $V = v$, the statistic S follows a multinomial distribution with parameters $(T+1, q_0(v), \dots, q_T(v))$. If this distribution were unconditional, the probabilities $q_0, \dots, q_T$ would be estimated by maximum likelihood, maximizing the log-likelihood: $\max_{q_0,\dots,q_T} \sum_{k=0}^{T} \frac{N_k}{N} \log(q_k)$, under the constraint that $\sum_{k=0}^{T} q_k = 1$, where N is the total number of individuals and N_k is the number of individuals for which $S = k$. This is close to what we do. In our case, N_k/N has to be replaced with the nonparametric estimate $\hat{q}_k(v)$. The unconstrained estimator of $q(v)$ maximizes $\max_{q_0,\dots,q_T} \sum_{k=0}^{T} \hat{q}_k(v) \log(q_k)$ under the constraint that $\sum_{k=0}^{T} q_k = 1$. The solution is $q_k(v) = \hat{q}_k(v)$, of course. We can now add the restriction that $(q_0(v), \dots, q_T(v))$ has to be such that $m(v) \in \mathcal{B}$. Using equation 19, for $k = 0, \dots, T$, we have $q_k(v) = c_0(v)C_k(v) \int u^k(1-u)^{T-k} dF(u|v)$. Since the unconstrained $\hat{m}(v)$ does not belong to $\mathcal{B}$, the constrained $\hat{m}(v)$ will belong to the border of $\mathcal{B}$. From theorem 3.4 (1), we know

that $F(.|v)$ has at most $T/2$ support points if T is even and at most $T/2+1$ support points if T is odd, including 0 or 1.

For the T even case, for instance, we thus solve the following problem:

$$\max_{p_0(v),\dots,p_{T/2}(v),w_1(v),\dots,w_{T/2}(v)} \sum_{k=0}^{T} \hat{h}_k(v) \log\left(c_0(v)C_k(v)\sum_{j=1}^{T/2} p_j(v)^k(1-p_j(v))^{T-k}w_j(v)\right)$$

$$\text{subject to } \begin{cases} p_1(v),\dots,p_{T/2}(v) \in [0,1] \\ w_1(v),\dots,w_{T/2}(v) \in [0,1] \text{ and } \sum_{j=1}^{T/2} w_j = 1 \\ \sum_{k=0}^{T} c_0(v)C_k(v)\sum_{j=1}^{T/2} p_j(v)^k(1-p_j(v))^{T-k}w_j(v) = 1 \end{cases}$$

Solving this problem directly gives us the unique principal representation. Note, however, that this method can prove computationally costly for big sample sizes because it involves many constrained optimization problems, with non linear constraints.

3.3.3 A second idea

With just two periods, the procedure we follow when $m(v) \in \mathcal{B}$ remains available when $m(v) \notin \mathcal{B}$. When $m(v) \in \mathcal{B}$, the sharp bounds for $\Delta(v)$ can directly be expressed using the $q_k(v)$, $k \in \{0,1,2\}$. The bound obtained with the upper principal representation reads $\overline{\Delta}(v) = q_1(v)\left(\frac{v_2}{v_1+v_2}\right)\left(\frac{q_1(v)v_1+q_0(v)(v_1+v_2)}{q_1(v)v_2+q_0(v)(v_1+v_2)}\right)$ and the bound obtained with the lower principal representation reads $\underline{\Delta}(v) = q_1(v)\left(\frac{v_1}{v_1+v_2}\right)\left(\frac{q_1(v)v_2+q_2(v)(v_1+v_2)}{q_1(v)v_1+q_2(v)(v_1+v_2)}\right)$. Provided that $\hat{q}_k(v)$, $k=0,1,2$ are all positive,[48] those quantities can always be computed even when the corresponding $m(v)$ does not belong to $\mathcal{B}$.

This approach has two advantages. First, it makes the estimator a smoother function of the data, which can help performing inference. Second, it is computationally costless. Its drawback is that it does not make much sense mathematically. When $m(v) \notin \mathcal{B}$, $\overline{\Delta}(v)$ and $\underline{\Delta}(v)$ have no more meaning. As a result this estimator overestimates even more the size of the identification region. The further $\hat{m}(v)$ away from $\mathcal{B}$, the wider the estimated identification region. Yet, the further $\hat{m}(v)$ away from $\mathcal{B}$, the likelier it is that $m(v)$ be close to the border of $\mathcal{B}$, the thinner the size of the real identification region.

[48] Even though those quantities are estimated probabilities, they can sometimes be negative depending on which nonparametric estimator is used.

3.3.4 More than two periods are available

When $T > 2$ periods are available, the first estimator can always be computed whereas the second one is unavailable. Yet, using the first one as described above might not be the best thing to do. Thanks to the use of an index, the dimension of the regressor in the nonparametric regression is T instead of $T \times K$. This is a substantial gain but T alone can still be too big. The solution we propose is to use the periods two by two. More precisely, using the periods k and T only, $k = 1, \ldots, T-1$, we can estimate bounds $l(X_k, X_T)$ and $u(X_k, X_T)$. We can then use all the partial bounds to perform inference on the true bounds for Δ. The average marginal effect can be written $\Delta = \mathbb{E}\left[\mathbb{E}\left[\Lambda'\left(X_T\beta_0 + \alpha\right)|X_1, \ldots, X_T\right]\right] \equiv \mathbb{E}\left[f\left(X_1, \ldots, X_T\right)\right]$. When X_k and X_T only are used, we estimate bounds $l(X_k, X_T)$ and $u(X_k, X_T)$ such that

$$l\left(X_k, X_T\right) \leq \mathbb{E}\left[f\left(X_1, \ldots, X_T\right)|X_k, X_T\right] \leq u\left(X_k, X_T\right)$$

Thus, the parameter Δ satisfies the $2T-2$ moment inequalities:

$$\begin{cases} \mathbb{E}\left[u\left(X_k, X_T\right) - \Delta\right] & \geq & 0 \\ \mathbb{E}\left[\Delta - l\left(X_k, X_T\right)\right] & \geq & 0 \end{cases}$$

We can then perform inference on Δ resorting to generalized moment selection, as described in Andrews and Soares (2010), provided we can estimate the standard errors for each conditional bound.

So far, we have unfortunately not been able to prove the consistency, let alone the asymptotic normality of our estimator. We thus try to convince readers our estimator works fine through simulations.

3.4 Monte-Carlo simulations

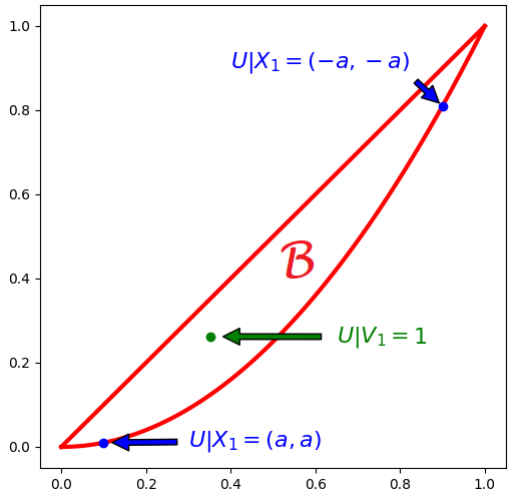
To fix ideas we start with a very simple data generating process. There are two periods and X is bi-dimensional. Let $X_{t,k}$ denote the value of the k-th covariate of X at period t. $X_{1,1}$ and $X_{1,2}$ can both take the value $a \equiv \log(9)$ with probability $1/2$ and the value $-a$ with probability 1/2. $X_{2,1}$ is always equal to 1 and $X_{2,2}$ is always equal to 0. Finally, we have $\beta_0 = (1 \quad -1)'$ and $\alpha = -X_{1,1}$.

For all $x_1 \in \{(-a,-a);(-a,a);(a,-a);(a,a)\}$, the distribution of α given X is a Dirac distribution. The corresponding moment vector admits a representation of index $1 =$

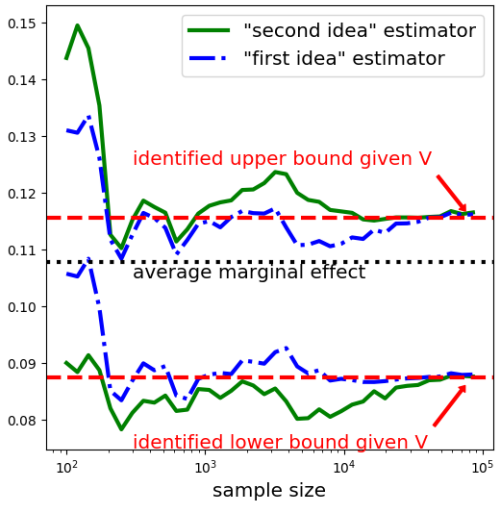
$T/2$, so theorem 3.4 tells us that this representation is unique. As a result, the average marginal effect is identified since each conditional marginal effect is. In this example, identification is lost when conditioning by V instead of X. Indeed, individuals for whom $v_1 = 1$ can either have $x_1 = (-a, -a)$ or $x_1 = (a, a)$. So for those individulas, the distribution of α is no more a Dirac distribution, it has now two points of support. The moment vectors are represented on figure 22 below. In this simple case we do not have to perform any nonparametric estimation. We see that our two different estimators are consistent.

Figure 22: A simple example

Identification



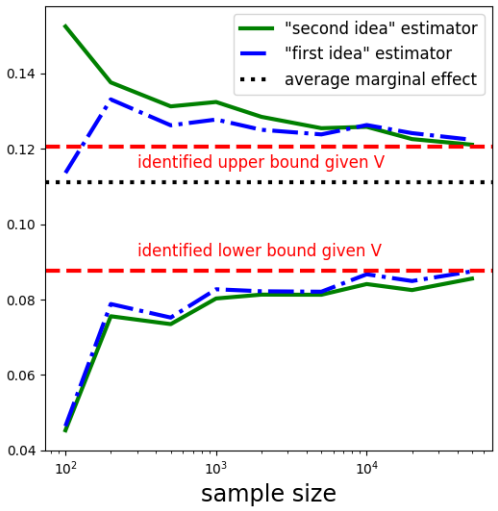
Consistency



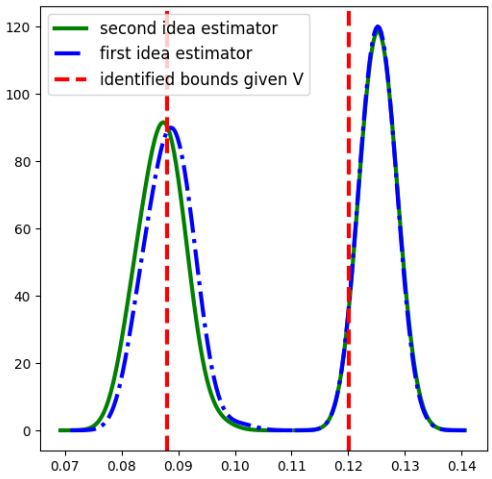
We now consider data generating process for which V is continuous. We start with $T = 2$: $X_{1,1}, X_{1,2} \overset{\perp\!\!\!\perp}{\sim} \mathcal{U}[-3, 3]$, $X_{21}, X_{2,2} \overset{\perp\!\!\!\perp}{\sim} \mathcal{U}[-2, 4]$. $X_1 \perp\!\!\!\perp X_2$, $\alpha = -X_{1,1}$ and $\beta_0 = (1 \quad -1)'$. This data generating process is simple enough for us to compute the exact true effect and the exact bounds, conditional to V. Figure 23 shows the consistency and the asymptotic normality of the estimators. For the asymptotic normality part, each estimate is computed with a sample of size $10,000$. We still see a small finite distance bias for the estimator of the upper bound.

Figure 23: $T = 2$

Consistency

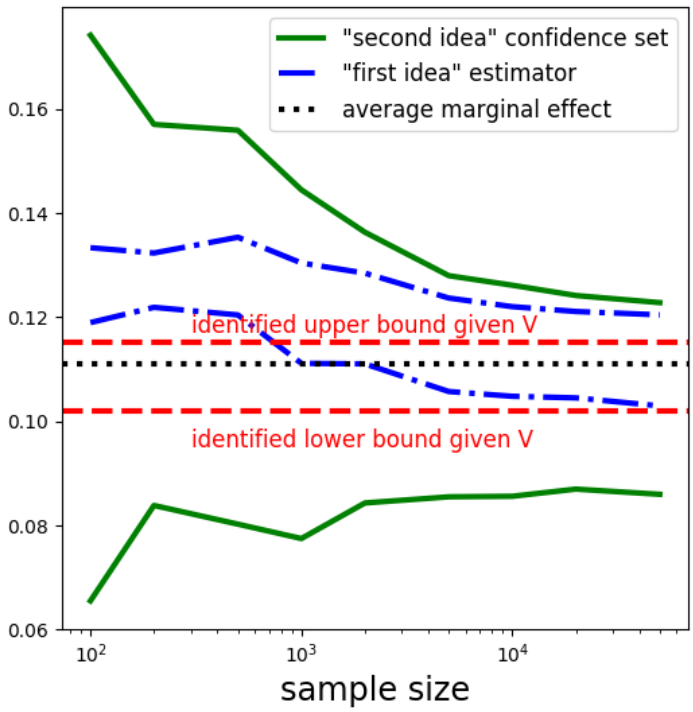


Asymptotic normality



Finally, we consider a very similar data generating process with $T = 3$. We have $\overset{\perp\!\!\!\perp}{\sim}\mathcal{U}[-3,3]$, $X_{21}, X_{2,2} \overset{\perp\!\!\!\perp}{\sim}\mathcal{U}[-2,4]$, $X_{31}, X_{3,2} \overset{\perp\!\!\!\perp}{\sim}\mathcal{U}[-4,2]$. X_1, X_2 and X_3 are mutually independent. $\alpha = -X_{1,1}$ and $\beta_0 = (1 \quad -1)'$. Figure 24 shows how our two estimators perform for this data generating process. For our first estimator, we display, like before, estimates of the identification bounds. For our second estimator, we display confidence intervals for the true parameter because this is what we directly obtain applying Andrews and Soares (2010)'s method. To implement this method, we need to provide a variance-covariance matrix for the four bounds estimated using two periods only. Since we have no formula for it, this matrix is estimated as a first step using simulations.

Figure 24: $T = 3$



The first idea estimator seems to be consistent and the second idea confidence intervals seem reasonable as well, although it is very difficult to assess their accuracy.

3.5 Conclusion

We have shown that the theory of Chebyshev systems could be used to change a difficult maximization problem into a very simple one and yield a nice characterization of the sharp bounds for the average marginal effect. We presented the method in a general enough way to show that it can be applied for other effects, which translate into other objective functions for the maximization problem. Quantile effects, for instance, are good candidates. This method can also certainly be used for totally different problems, as long as the Chebyshev assumption is satisfied. Investigating in which other cases this method can be applied is left for future research.

General conclusion

Bitcoin

Mr. Julien Prat and I provide the first dynamic equilibrium model of the market for Bitcoin mining. This work provides both a successful empirical test of the theoretical model developed in Caballero and Pyndick (1996), and a deeper understanding of bitcoin miner's behavior. I then rely on this new insight to suggest a simple solution to lower Bitcoin's electricity consumption. Both chapters also highlight the preponderant role played by the technical progress rate of mining hardware. The lower this rate, the higher the aggregate level of investment, the higher the network electricity consumption and the less that consumption can be cut using the solution I advocate. If the idea proposed in the second chapter is to be implemented, waiting is not only a loss of time, it is detrimental since the technical progress rate keeps decreasing.

These two chapters have the following limitations. The model developed in the first chapter can predict the hashrate relatively accurately on the medium and long term. Yet, it fails to do it well in the short term, mainly due to manufacturing / delivery delays which are not accounted for. An obvious direction for future research would be to try and improve the model this way. For the sake of simplicity, the second chapter does not exactly build on the first chapter's model but on a simplified version of it, where the ₿/ $ exchange rate is deterministic. If such an idealized model remains realistic enough to convey the main idea of the suggested protocol change, a more realistic model may be needed to accurately calibrate the different parameters and make sure the protocol change will not harm active miners. Future research should then try adapting this toy model to the exact framework developed in the first chapter.

Panel data and binary dependent variables

In the third chapter, Messrs. Davezies and D'Haultfœuille and I provide a new and satisfying method to bound the average marginal effect of explanatory variables, one of the main quantity of interest, when the dependent variable is binary and when panel data are available. I deem this method as satisfying for two reasons. First, very few, if any, strong and embarrassing assumptions have to be made, as opposed to the linear or the random effect logit models. Second, our estimator can be computed very

quickly and can thus be applied on pretty massive data sets. The framework in which our proposed estimator is available is very broad and so I believe that this new method will be of interest for many applied researchers.

The main drawback of this chapter is obvious: neither the consistency nor asymptotic normality, let alone the uniformly valid inference property of our estimator are proved so far. Finding under which exact set of technical assumptions those results hold and proving them should be the priority for future research. As a byproduct, the method developed in this chapter brought to the field of economics a not very-well-known mathematical result. Mathematical problems similar to the one we solved can probably be encountered in many other frameworks. Other effects, such as quantile effects, for the same model are good candidates. Finding out other situations where Chebyshev systems results can be applied is also a promising future research track.

References

Aid, R., Federico, S., Pham, H., and Villeneuve, B. (2015). Explicit Investment Rules with Time-to-Build and Uncertainty. *Journal of Economics Dynamics and Control*, pages 240–256.

Alvarez, F. and Shimer, R. (2011). Search and rest unemployment. *Econometrica*, 79:75–122.

Andrews, D. W. K. and Soares, G. (2010). Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica*, 78:119–157.

Antonopoulos, A. M. (2014). *Mastering Bitcoin*. O'Reilly'.

Assia, Y., Buterin, V., m, l., Rosenfeld, M., and Lev, R. (2013). Coloured coins White paper. https://docs.google.com/document/d/1AnkP_cVZTCMLIzw4DvsW6M8Q2JC0lIzrTLuoWu2z1BE/edit.

Athey, S., Parashkevov, I., Sarukkai, V., and Xia, J. (2017). Bitcoin Pricing, Adoption, and Usage: Theory and Evidence. Stanford University Graduate School of Business Research Paper No. 16-42. Available at SSRN: https://ssrn.com/abstract=2826674.

Back, A., Corallo, M., Dashjr, L., Friedenbach, M., Maxwell, G., Miller, A., Poelstra, A., Timon, J., and Wuille, P. (2014). Enabling Blockchain Inovations with Pegged Sidechains. https://blockstream.com/sidechains.pdf.

Bertola, G. and Caballero, R. (1994). Irreversibility and aggregate investment. *The Review of Economic Studies*, 61:223–246.

Biais, B., Bisière, C., Bouvard, M., and Casamatta, C. (2017). The blockchain folk theorem. TSE Working paper n. 17-817,https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/wp/2017/wp_tse_817.pdf.

Bontemps, C., Magnac, T., and Maurin, E. (2012). Set identified linear models. *Econometrica*, 3:1129–1155.

Bowden, R., Keeler, H., Krezinski, A., and Taylor, P. (2018). Block arrivals in the bitcoin blockchain. https://arxiv.org/pdf/1801.07447.pdf.

Caballero, R. J. and Pyndick, R. S. (1996). Uncertainty, Investment, and Industry Evolution. *International Economic Review*, pages 641–662.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies*, 47(1):225–238.

Chamberlain, G. (2010). Binary response models for panel data: Identification and information. *Econometrica*, 78:159–168.

Chernozhukov, V., Fernandez-Val, I., Hahn, J., and Newey, W. (2013). Average and quantile effects in nonseparable panel models. *Econometrica*, 81(2):535–580.

Chesher, A., Rosen, A. M., and Smolinski, K. (2013). An instrumental variable model of multiple discrete choice. *Quantitative Economics*, 4:157–196.

Chiu, J. and Koeppl, T. (2017). The Economics of Cryptocurrencies - Bitcoin and Beyond. Available at SSRN: https://ssrn.com/abstract=3048124.

Cong, L. W. and He, Z. (2017). Blockchain disruption and smart contracts. https://philadelphiafed.org/-/media/bank-resources/supervision-and-regulation/events/2017/fintech/resources/blockchain-disruption-smart-contracts.pdf?la=en.

Davezies, L. and D'Haultfœuille, X. (2016). A new characterization of identified sets in partially identified models. http://www.crest.fr/ckfinder/userfiles/files/Pageperso/ldavezies/Work-In-Progress/paper_2016.pdf.

Decker, C. and Wattenhoffer, R. (2013). Information Propagation in the Bitcoin Network. *13-th IEEE International conference on peer-to-peer computing*. https://github.com/bellaj/Blockchain/blob/master/Information%20Propagation%20in%20the%20Bitcoin%20Network.pdf.

D'Haultfœuille, X. and Rathelot, R. (2011). Measuring segregation on small units: A partial identification analysis. *Quantitative Economics*, 8.

Dixit, A. K. and Pyndick, R. S. (1994). *Investment under Uncertainty*. Princeton University Press.

Fernández-Villaverde, J. and Sanches, D. (2016). Can Currency Competition Work? http://economics.sas.upenn.edu/~jesusfv/currency_competition.pdf.

Foley, S., Karlsen, J. R., and Puntnis, T. J. (2018). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3102645.

Galichon, A. and Henry, M. (2011). Set identification in models with multiple equilibria. *Review of economic studies*, 78:1264–1298.

Gamboa, F. and Lozada-Chang, L.-V. (2004). Large deviations from random power moment problem. *The Annals of Probability*, 32:2819–2837.

Gandal, N., Hamrick, J., Moore, t., and Oberman, T. (2017). Price Manipulation in the Bitcoin Ecosystem. http://weis2017.econinfosec.org/wp-content/uploads/sites/3/2017/05/WEIS_2017_paper_21.pdf, CEPR Working paper.

Grunspan, C. and Pérez-Marco, R. (2017). Double spend races. https://arxiv.org/pdf/1702.02867.pdf.

Harrison, M. J. (2013). *Brownian Models of Performance and Control*. Cambridge University Press.

Hoderlein, S. and White, H. (2012). Nonparametric identification in nonseparable panel data models with generalized fixed effects. *Journal of Econometrics*, 168(2):300–314.

Hong, K., Park, K., and Yu, J. (2017). Crowding out in a Dual Currency Regime? Digital versus Fiat Currency. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2962770, Bank of korea working paper.

Houy, N. (2016). The Bitcoin Mining Game. *LEDGER*.

Huberman, G., Leshno, J. D., and Moallemi, C. C. (2017). Monopoly without a monopolist: An economic analysis of the bitcoin payment system. https://moallemi.com/ciamac/papers/bitcoin-2017.pdf.

Kaido, H. and Santos, A. (2014). Asymptotically efficient estimation of models defined by convex moment restrictions. *Econometrica*, 82:387–413.

Karame, G. O., Androulaki, E., and Capkun, S. (2012). Two Bitcoins at the Price of one? Double-spending Attacks on Fast Payments in Bitcoin. https://eprint.iacr.org/2012/248.pdf.

King, S. (2013). Primecoin: Cryptocurrency with prime number proof-of-work. http://primecoin.io/bin/primecoin-paper.pdf.

King, S. and Nadal, S. (2012). Ppcoin: Peer-to-peer crypto-currency with proof-of-stake. https://peercoin.net/assets/paper/peercoin-paper.pdf.

Krein, M. G. and Nudel'man, A. A. (1977). *The Markov Moment Problem and Extremal Problems*. Translations of Mathematical monographs.

Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of econometrics*, 95:391–413.

Lerner, S. (2015). Rsk: Bitcoin powered Smart Contracts. http://www.the-blockchain.com/docs/Rootstock-WhitePaper-Overview.pdf.

Ma, J., Gans, J. S., and Rabee, T. (2018). Market structure in bitcoin mining. http://www.nber.org/papers/w24242?utm_campaign=ntw&utm_medium=email&utm_source=ntw.

Manski, C. F. (2003). *Partial identification of probability distributions*. Springer series in statistics.

Nakamoto, S. (2008). Bitcoin, a Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf.

Poon, J. and Thaddeus, D. (2016). The Bitcon Lightning Network: Scalabble Off-chain Instant Payments. https://lightning.network/lightning-network-paper.pdf.

Popov, S. (2017). The tangle. https://iota.org/IOTA_Whitepaper.pdf.

Prat, J. and Walter, B. (2018). An equilibrium model of the market for bitcoin mining. CESifo Working paper n. 6865,https://www.cesifo-group.de/DocDL/cesifo1_wp6865.pdf.

Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Medicine*, pages 321–333.

Reid, F. and Harrigan, M. (2012). An Analysis of Anonimity in the Bitcoin System. arXiv:1107.4524v2.

Rosenfeld, M. (2011). Analysis of Bitcoin Pooled Mining Reward Systems. https://arxiv.org/pdf/1112.4980.pdf, arXiv:1112.4980.

Schilling, L. and Uhlig, H. (2018). Some simple bitcoin economics. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3155310.

Szabo, N. (1996). Building Blocks for Digital Markets. http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_2.html.

université PARIS-SACLAY | ÉCOLE DOCTORALE Sciences de l'homme et de la société (SHS)

Titre : Deux essais sur le marché des mineurs de bitcoins et un essai sur le modèle logit avec effets fixes et données de panel.

Mots clés : bitcoin, systèmes transactionnels, investissements, économies d'énergie, modèles de choix discrets, panels.

Résumé : Ma thèse se compose de deux parties indépendantes. La première traite de crypto-économie et la seconde d'économétrie théorique. Dans le premier chapitre, je présente un modèle qui prédit la puissance de calcul totale déployée par les mineurs en utilisant le taux de change bitcoin / dollar. Le deuxième chapitre s'appuie sur une version simplifiée du précédent modèle pour faire le constat de l'inefficacité du protocole Bitcoin actuel et proposer un moyen simple de réduire la consommation d'électricité engendrée par cette cryptomonnaie. Le troisième chapitre explique comment identifier et estimer les bornes exactes de la région d'identification de l'effet marginal moyen dans un modèle logit avec effets fixes sur données de panel.

Title : Two essays on the market for bitcoin mining and one essay on the fixed effects logit model with panel data.

Keywords : bitcoin, payment systems, investments, energy savings, discrete choice models, panels.

Abstract : My dissertation concatenates two independent parts. The first one deals with crypto-economics whereas the second one is about theoretical econometrics. In the first chapter, I present a model which predicts bitcoin miners' total computing power using the bitcoin / dollar exchange rate. The second chapter builds on a simplified version of the preceeding model to show to which extent the current Bitcoin protocol is inefficient and suggest a simple solution to lower the cryptocurrency's electricity consumption. The third chapter explains how to identify and estimate the sharp bounds of the average marginal effect's identification region in a fixed effects logit model with panel data.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France