

Analysis of ecrime in Crowd-sourced Labor Markets: Mechanical Turk vs. Freelancer

Vaibhav Garg, Chris Kanich, L. Jean Camp

Abstract

Research in the economics of security has contributed more than a decade of empirical findings to the understanding of the microeconomics of (in)security, privacy, and ecrime. Here we build on insights from previous macro-level research on crime, and microeconomic analyses of ecrime to develop a set of hypotheses to predict which variables are correlated with national participation levels in crowd-sourced ecrime. Some hypotheses appear to hold, e.g. Internet penetration, English literacy, size of the labor market, and government policy all are significant indicators of crowd-sourced ecrime market participation. Greater governmental transparency, less corruption, and more consistent rule of law lower the participation rate in ecrime. Other results are counter-intuitive. GDP per person is not significant, and unusually for crime, a greater percentage of women does not correlate to decreased crime. One finding relevant to policymaking is that deterring bidders in crowd-sourced labor markets is an ineffective approach to decreasing demand and in turn market size.

1 Introduction

The new school of ecrime [50] is both organized and driven by profits [10, 35, 52]. This is reflected by the technical and policy proposals to fight ecrime, all grounded in deterrence or rational choice theory [43]. As Anderson demonstrated in his canonical work, technical measures will never be the silver bullet [2]. Policy efforts must complement technical measures [30]. Thus, both security researchers as well as practitioners must address the economic incentives (or disincentives) to invest in security by different stakeholders [50]. To the extent that security is a market, there are legitimate stakeholders such as service providers and end-users. Economic analysis of the incentives that drive service providers [18] and end-users [27] have offered academic insights [41] that inform practical solutions, e.g. Google Vulnerability Reward program.

International efforts been narrowly focused on ecrime itself, e.g., the European Convention of Cybercrime. A macro-level approach offers a set of complementary tactics for decreasing the threat; that is, potential measures to reduce the motivations for attackers [28]. A broader response focuses on creating environments where ecrime would be unlikely to flourish. This would go beyond the immediate term effort to deter individual attackers; instead it seeks to build a long term structure so that attackers to become legitimate market participants.

To the extent that all market participants are drive by profits, why do some choose to become legitimate stakeholders, while other resort to criminal activity? Often we observe that criminal activity online clusters in specific countries, e.g. Romania, Nigeria, indicates that such determination is informed by at least in part by macroeconomic factors. In previous research, we developed a macroeconomic model of organized ecrime [19] based on economics of smuggling [4]. We concluded

that organized crime can be welfare increasing in local jurisdictions. We found that crime markets may exist in one of two possible equilibria: 1) high enforcement low crime or 2) low enforcement high crime. Despite the near term increase in social welfare, a low enforcement high crime equilibrium is not ideal, as thriving crime potentially acts as a prohibitive tariff against a legitimate market.

In this paper, we empirically examine the theoretical findings of our previous work. We use the geographic locations of crime crowd-sourced labor to evaluate existence of nations in the dual equilibria. Specifically, we identify the macro-level variables that encourage participation in legitimate crowd sourcing markets, and distinguish them from variables that appear to facilitate illegitimate crime activities. Section 2 is background and related work. Methodology is described in section 3. Section 4 details the results. Section 5 presents the discussion. We conclude in Section 6.

2 Background & Related Work

Anderson notes the economic nature of security markets [2]. Both attackers [10, 35, 52] and defenders [55] are economically incentivized. Franklin et al. [17] note the shift from ‘hacking for fun’ to ‘hacking for profit’. Participants have become specialized to improve efficiency [36]. The goods being traded range from zero day vulnerabilities [34] to human CAPTCHA solvers [37]. Increasing activity in underground markets of information goods to bypass security measures has led to financial loss for individuals and institutions alike. The annual loss due to phishing, and possible gain to phishers, has been claimed to be as big as \$178.1 million dollars a year [35].

Motoyama et al. [38] examine one such market in Freelancer. Freelancer is a crowd-sourced labor market. Participants in the market are either bidders or buyers. Buyers create demand by posting jobs that are difficult to automate, but relatively easy to do for humans, e.g. transcription, translation. Bidders bid on the jobs of their choice and are compensated based on their performance or other criteria stated by the buyer. However, only upto 65.4% of the jobs posted on Freelancer are for legitimate tasks. Several jobs ask the bidders to do tasks that thwart security mechanisms, e.g. solve CAPTHCAs, spam.

Web service abuse is not limited to crowd sourced labor markets, nor is the abuse limited to online crime. Thomas et al. [52] found that Internet Relay Chat (IRC) channels are being used to trade credit card data and other financial information. They also find evidence of physical crime. Holz et al. analyze the underground market with the instance of keyloggers and dropzones [21]. Threats with externalities, such as malware, may have a greater impact even when a subset of end-users are conscientious [29]. In fake antivirus, attackers have found a way of duping conscientious end-users who might not be technically adept. Stone-Gross et al. [51] estimated that the combination of merely three fake antivirus businesses generated a revenue of approximately \$130 million dollars.

Simultaneously, economics has also informed defender strategies to alleviate cybercrime. On the technical side, the goal has been to make attacks more expensive and decrease the rational imperative to attack through diminishing returns [30]. Such deterrence-based approaches are potentially successful [20], but the impact may be limited to small time frame [43].

Thus, many previous investigations have either been microeconomic or game theoretic, the former investigating attacker motivations and the latter suggesting defender strategies. Researchers have targeted specific markets such as IRC channels [52], Freelancer [38], malicious Chinese websites [61]. Complementary work using macroeconomics, however, has been limited.

However, insights grounded in macroeconomics are much needed. For example, the policy solutions to massive copyright violations based in deterrence theory have resulted in misguided regulatory proposals such as SOPA/PIPA. However, macroeconomic modeling of software copyright violation notes that violations in many instances would lead to an increase in revenue due to externalities [40]. Furthermore, Osorio found that massive copyright violations are driven by lack of access and economic resources [40]. Arguably, then Netflix has been more effective than SOPA would be [48]. Simultaneously, there is evidence that price cuts [9] have been more effective than DMCA [16].

Macro-level analysis has been used to study a diverse set of problems from smuggling [4] to olympic gold medals [3]. Bernard et al. [3] developed a regression based empirical model that was able to predict the number of olympic medals won by every country. Bhagwati et al. [4] on other hand made a theoretical model of smuggling. Counterintuitively, they found that smuggling is social welfare increasing. While smuggling denies trade gain, it engenders production and consumption gain. As long as the sum of production and consumption is greater than trade loss, smuggling would be social welfare increasing.

Macro-level investigations have been used to study other organized criminal activities offline. There are seven theories in criminology that have been empirically validated for crime in the physical world [45]. Social disorganization theory suggests that crime is a manifestation of neighborhood dynamics rather than that of individual motivations. Influential factors include urbanism, poverty, residential transience, heterogeneity as well as family disruption [47]. While investigations into this theory are relatively new, the findings have been encouraging [45]. Measuring informal social control can, however, be difficult especially over the Internet.

Anomie/Strain theory underlines the disconnect between culturally driven individual aspirations and the social structures that facilitate achievement. When opportunities are rare there is incentive for individuals to deviate from cultural norms to achieve culturally desirable goals. For example, Messner et al. [33] modeled crime as a function of American dream, which is driven by an emphasis on economic success and provided for by an institutional structure built on economy. While macro level assessment of this theory is rare, they measure the strength of non economic institutions. Some influential factors include family structure, religious participation, political involvement, education, and access to welfare, income-replacement value of welfare, as well as it's comprehensiveness [7, 49]. While there is significant support for this theory, directly measuring the strength of relevant macro indicators is difficult [45].

Resource/Economic deprivation theory analyses, both the impact of poverty as well as economic inequality, e.g. income disparity [59, 5]. Thus, deprivation can be relative or absolute. Both perspectives have been extensively tested and demonstrate a strong and reliable ability to impact crime [59, 45]. This theory has support in cybercrime, for example previous research has identified software piracy as a function of Gross Domestic Product (GDP) per capita¹ [40].

Routine activity theory assumes motivated offenders and examines the macro level indicators that engender opportunities for the offenders to exploit. Convergence of offender, target and absent guardianship drives deviant behavior. The key measures here are household activity ratio and aggregate unemployment. Empirical validation of this theory, with most studies concentrating on lack of guardianship or lack of informal social control [12]. The theory is very pertinent to cybercrime. Irrespective of the motivations, driven attackers or offenders are always around. Same is true for targets that are weakly protected. However, this theory has received little support in the

¹Gross Domestic Product (GDP) indicates the aggregate worth of goods and services produced by a country in a specific time frame, typically annually. GDP per capita is an indicator of the average standard of living in a country

domain of cybercrime [60].

Deterrence/Rational choice theory analyses the impact of deterrence initiatives on crime, e.g. incarceration, criminal justice system, regulation, prosecution etc. While the impact of this theory has weak empirical validation for crime in general [45], it has support for cybercrime [43, 44].

Social support/altruism theory looks at the inverse relationship between state sponsored support [13] or community altruism [8] with crime rates. This theory again has limited empirical validation. It is also difficult to demonstrate the difference between state sponsored support, e.g. welfare, vs. private altruism, e.g. charity.

Subcultural theory examines if certain cultures are predisposed towards deviant behavior. For example, Colin Powell famously called Nigeria ‘a nation of scammers’. The underlying determinants of predisposition towards crime may be other factors such as large urban population. This is possibly the weakest theory in terms of empirical support. To the extent that culture is reflected by legal frameworks, there is some support for this theory in massive copyright infringement [40].

A similar cohesive theory of ecrime is missing. Our ability to fight ecrime is limited by our understanding of the actors involved. While the microeconomic investigations have provided insight into the structure of ecrime markets and how they function [51, 38], they have been limited in their ability to explain the evolution of organized ecrime. Why are certain markets more conducive to deviant behavior than others? Which markets would be more conducive to what kind of ecrime? In this paper, we present a first such investigation for crowd-sourced markets.

3 Methodology & Data Collection

In previous research, we developed a theoretical macroeconomic model of ecrime [19]. Our model was based in the macroeconomic analysis of smuggling. We assumed that illegal goods are smuggled analogues of legal goods and are therefore perfectly substitutable. We found that organized ecrime can be profit increasing in local jurisdictions. Further, we argue that the success of illegal goods can act as a prohibitive tariff for the development of local legal markets. As such the market exists in one of two equilibria: 1) high enforcement low crime or 2) low enforcement high crime. Our findings would be generalizable to markets where legal and smuggled goods coexist [42].

In this paper, we conduct an empirical analysis of the macro-level factors that drive the market towards either equilibrium. We consider the specific instance of crowd-sourced labor markets. High enforcement and low crime is represented by Amazon’s Mechanical Turk. Low enforcement and high crime is represented by Freelancer. Both Mechanical Turk and Freelancer provide functionally equivalent services, i.e. the ability to crowd-source tasks that are difficult or expensive to automate through computers, but require relatively less effort for human agents.

Amazon’s Mechanical Turk service is used as an example of a high enforcement low crime crowd sourced market. While Mechanical Turk is usually used for legitimate purposes, such as for survey-based research by academics [31]. Simultaneously, Mechanical Turk can potentially be used for illegitimate activities, e.g. CAPTCHA solving [6], Malware installations [24, 11], . The demographic distribution of Mechanical Turk workers has been studied by Ross et al. [46] and Ipeirotis [23]. Both studies provide a demographic analysis of Turkers who participate by bidding for Human Intelligence Tasks (HITs). The analysis in this paper uses a country-based distribution of Mechanical Turk workers from the publicly available database by Ipeirotis².

²<http://hdl.handle.net/2451/29585>, Retrieved on 24 February 2012

Freelancer is an example of a market that has low enforcement and high crime. Motoyama et al. provide a distribution of two kinds of Freelancer participants: bidders and buyers [38]. Bidders are more akin to Mechanical Turk workers, in that they bid on tasks provided by other participants. Buyers, however, are the participants that are responsible for the market to exist as they create the demand by posting jobs that need to be completed. 65.4% of Freelancer jobs are legitimate. However, Freelancer similar to Mechanical Turk can and is used for solving CAPTCHAs [37]. Additionally, it is used for other undesirable activities, such as account creation, social networking link generation and search engine optimization support. Motoyama et. al [38] identified 22 distinct job types of Freelancer classified in six categories: 1) legitimate, 2) accounts, 3) search engine optimization (SEO), 4) spam, 5) OSN, and 6) miscellaneous. The analysis in this paper considers four of the identified categories:

1. Accounts: CAPTCHA solvers, basic accounts, and verified accounts.
2. SEO: white hat links, grey hat links, and miscellaneous.
3. Spam: bulk email, and bulk advertisement.
4. OSN: Social network links.

Note that we assume that Mechanical Turk and Freelancer are perfect substitutes, due to assumptions of the underlying theoretical model [19]. However, this assumption is made for simplicity and the results would be applicable even when the same individual participates in both markets. A second assumption is that Mechanical Turk market is primarily honest, while Freelancer is illegal. This too is an artifact of the underlying theoretical model [4, 19]. However, Pitt has shown that the results hold true even when such distinction is not clear, i.e. honest and illegal markets coexist [42].

We consider participation in either crowd-sourced markets, Mechanical Turk or Freelancer, as a function of several macro-level factors. Osorio showed that macroeconomic indicators such as GDP per capita are predictors of massive copyright infringement [40]. Osorio consider a three dimensional model: (1) accessibility, (2) affordability, and (3) legal framework. Accessibility was operationalized as the ability of the software to fit local needs, presence of after sales support and corporate presence. Affordability was operationalized as GDP per capita. Legal framework was operationalized using the work of Easterly et al. [14]. Osorio's paper empirically examined the theoretical assertions of prior research [1, 25, 54].

Similarly, we begin by considering the theories that have found support in Osorio's analysis of copyright infringement. A key determinant in Osorio's model was Gross Domestic Product (GDP) per capita. The wages afforded to either Mechanical Turk or Freelancer workers are much lower than the minimum wage requirements in USA [22]. This indicates that there is an economic imperative to participate. Mason et al.[32] found that financial incentives do increase the quantity of participation for Mechanical Turk. While the wages are low, the corresponding value in local markets might be high based on purchasing power parity (PPP). Thus, another variable to consider would be GDP per capita by Purchasing Power Parity (PPP)³. Both GDP per capita and GDP per capita by PPP are available from World Bank Development Indicators (WDI) ⁴.

³Ideally, identical goods cost the same in two different markets, when priced in the same currency. However, transaction costs lead to different prices. Purchasing Power Parity measures the difference between prices in two different markets for identical goods and services.

⁴<http://data.worldbank.org/indicator>, Retrieved on 24 February 2012

Accessibility and affordability were the other two measures in Osorio’s model. These are driven by the extant conditions of local ICT markets. To the extent that ICT investment, both public and private, is available, it would make participation in either Mechanical Turk or Freelancer easier. We operationalize this by using Digital Economy Rankings, produced by Economist Intelligence Unit in collaboration IBM Institute for Business Value [53]. These rankings capture more than eReadiness of the country, evaluating quality as well as quantity. For example, they measure Internet penetration as well as speed, connectivity, affordability etc. The Digital Economy Rankings are a linear combination of six factors:

1. Connectivity and technology infrastructure: This indicates access to affordable connectivity, for both broadband and mobile, measuring assurance quality, reliability, and security.
2. Business environment: This indicates the degree to which development in private sector is facilitated by economy, political stability, taxation, competition policy, the labour market, and openness to trade and investment
3. Social and cultural environment: This measures both formal education as well as Internet literacy and associated technical skills.
4. Legal environment: This quantifies the progressive nature of local legislative framework, pertaining to Internet commerce, to combat ecrime, spam etc. as well as abuses and non-competitive behavior.
5. Government policy and vision: Technology adoption by the government to facilitate citizen participation as well and access to information.
6. Consumer and business adoption: Adoption of existing digital channels by businesses and individuals.

We also considered export of Information and Communication Technology (ICT) services as well as percentage export of ICT services from WDI. These measure the net worth of the ICT goods exported, software is excluded. Percentage is computed as a ratio with net worth of all goods exported. To a degree this measures the success of the business environment evaluated in the digital economy rankings. Similarly we complement legal environment with *rule of law* from Worldwide Governance Indicators (WGI) [26]. Rule of law indicates the degree to which legal framework is implemented. Legal framework can also be thwarted by corruption or perceptions thereof. The former is measured by the corruption index from Transparency International (TI) as well as *control of corruption from WGI*. TI’s corruption index as well as WGI’s *control of corruption* measure perceptions of corruption, where corruption is defined as misuse of public power for private gain.

Other measures from WGI are also found to complement digital economy rankings. Both *government effectiveness* and *regulatory quality* are considered along with government policy and vision. Government effectiveness measures the perceived quality of public services, quality of civil service and the degree to which it is independent from political manipulation, the quality of policy formulation and implementation, and the perceived credibility of the government to commit to said policies. Regulatory quality quantifies the perceived ability of the government towards sound policy/regulations formulation and implementation that encourage private sector development. *Voice and accountability* is complementary to social and cultural environment. It measures intellectual and political freedom.

We also consider additional WDI indicators related to availability of labor, measured by: 1) population, 2) population percentage of women, 3) percentage of urban population, and 4) number of Internet users. Population and number of Internet users is an indicator of the available labor pool. Simultaneously, participation in crowd-sourced labor markets is made possible by Internet adoption. However, gender based differences in adoption preferences may shift the equilibrium towards mechanical turk or vice versa [56]. Women are also less likely to commit crime offline [39, 15]. Thus, higher ratio of women may shift the equilibrium toward Mechanical Turk. Urban are more likely to have better access to technological infrastructure. Thus, higher proportion of urban population would lead to higher Internet adoption, and therefore higher participation in crowd sourced markets. Its impact on market equilibria would be insightful. Both percentage of women as well as percentage of urban population are available from WDI.

We consider language proficiency skills, specifically English language proficiency, as another macro indicator. English language proficiency is different from formal education as measured by social and cultural factors. Legitimate crowd-sourced tasks such as survey participation or proof-reading require a degree of fluency with the specific language, mostly English as most tasks on Amazon are in English. However, illegal tasks such as CAPTCHA solving at best requires a mechanical pattern recognition that comes easy to human agents [57] than automated ones [58]. For Mechanical Turk, a minimum level of English proficiency would be required to be able to understand the job solicitations, which are typically made in English as requests can only be submitted from United States. English language proficiency is operationalized using the TOEFL’s ranking of countries on reading, speaking, listening, and writing⁵.

Finally, Security of ICT infrastructure is also an indicator: 1) number of secure Internet servers (SIS), and 2) number of SIS by population. SIS and SIS by population would encourage market investment by providing assurance of security. These indicators have also been procured from WDI. The final regression equation is given by equation 1, where N corresponds to number of workers; AFF, ACC, LEG, POP, ENG, and SEC refer to measures of affordability, accessibility, legal framework, availability of labor, English language proficiency, and security respectively. A list of all variables considered and respective sources is given in Table 1.

$$N = \beta_0 + \beta_1 * AFF + \beta_2 * ACC + \beta_3 * LEG + \beta_4 * POP + \beta_5 * ENG + \beta_6 * SEC \quad (1)$$

4 Results

In this paper, we empirically examine a macro-economic model of organized ecrime by considering the specific example of crowd-sourced labor markets. The model posited that information communication technology markets would tend to exist in one of two equilibria: 1) high enforcement low crime, and 2) low enforcement high crime [19]. Amazon’s Mechanical Turk is an instance of high enforcement low crime, while Freelancer represents low enforcement high crime. A list of all the countries considered is given in Appendix A.

Participation in either of these markets is a function of macro-level indicators, including macroeconomic indicators. The research question that we address is here is two fold. Which macro-level indicators encourage participation in crowd-sourced labor markets? Secondly, which specific indicators inform the market equilibrium, by either encouraging or alleviating criminal activity online.

⁵https://www.ets.org/toefl/research/topics/candidates_and_populations, Retrieved on 24 February 2012

Table 1: List of Macro-Level Variables

Affordability (AFF)		
GDP per capita	WDI	2010
” by PPP	WDI	2010
Consumer & business adoption	Economist	2010
Accessibility (ACC)		
Digital Economy Rankings	Economist	2010
Connectivity & Technology	Economist	2010
Business Environment	Economist	2010
Export of ICT services	WDI	2010
% export of ICT services	WDI	2010
Social & Cultural Environment	Economist	2010
Voice & Accountability	WGI	2010
Legal (LEG)		
Legal Environment	Economist	2010
Rule of law	WGI	2010
TI corruption index	Transparency International	2011
Control of corruption	WGI	2010
Government policy & vision	Economist	2010
Government effectiveness	WGI	2010
Regulatory quality	WGI	2010
Population (POP)		
Population	WDI	2010
Population density	WDI	2010
Population % of women	WDI	2010
% of urban population	WDI	2010
Number of Internet users	WDI	2009
English (ENG)		
English reading	TOEFL	2010
English listening	TOEFL	2010
English speaking	TOEFL	2010
English writing	TOEFL	2010
Security (SEC)		
SIS	WDI	2010
SIS by population	WDI	2010

Table 2: Best Fit Models & Corresponding p-values

Macro-level Indicator	Mechanical Turk	Freelancer Bidder	Freelancer Buyer
Affordability (AFF)			
GDP per capita ” by PPP per capita Consumer & business adoption	0.05		0.05
Accessibility (ACC)			
Digital economy ranking	0.05		0.01
Connectivity & technology	0.01	NS	0.001
Business environment			0.01
Export of ICT services		NS	
% export of ICT services		NS	
Social & cultural environment	NS	0.01	0.05
Voice and accountability	NS		NS
Legal (LEG)			
Legal environment	NS		
Rule of law		0.05	NS
TI corruption Index			NS
Control of corruption		NS	0.05
Government policy & vision	NS	NS	0.05
Government Effectiveness		0.05	
Regulatory quality			NS
Population (Pop)			
Population	0.001	0.001	0.001
Population % of women	NS		0.05
% of urban population			0.05
Number of Internet users	0.001	0.001	0.001
English (ENG)			
English reading	NS	NS	NS
English listening	0.01		0.01
English speaking	0.05		
English writing	NS	NS	
Security (SEC)			
SIS	0.001	0.001	0.001
SIS by population	0.01	NS	NS
Adjusted R-square's p-value<2.2e-16			
Adjusted R-square	0.9706	0.9481	0.9625

Table 3: Exploratory Factor Analysis for Mechanical Turk

Mechanical Turk	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
Digital economy ranking	0.973	0.220			0.005
Connectivity & technology	0.934	0.138		0.187	0.067
Social & cultural environment	0.930	0.196		0.108	0.080
Legal environment	0.823	0.414		-0.212	0.105
Government policy & vision	0.938	0.193		-0.174	0.053
Consumer & business adoption	0.959	0.221			0.028
English reading	0.294	0.860			0.173
English listening	0.347	0.900	-0.177	0.180	0.005
English speaking	0.287	0.768	-0.219	0.304	0.188
English writing	0.303	0.864			0.161
Number of Internet users		-0.133	0.988		0.005
Population	-0.278		0.778		0.312
Population % of women		0.404		-0.107	0.825
Voice and accountability	0.730	0.463	-0.112		0.235
Secure Internet servers (SIS)	0.384		0.490		0.593
SIS by population	0.760	0.201		0.322	0.278
Proportion Variance	0.43	0.231	0.121	0.024	
Cumulative Variance	0.43	0.661	0.782	0.805	

In section 3, we operationalized these indicators for a six dimensional framework: 1) affordability, 2) accessibility, 3) legal framework, 4) availability of labor, 5) English language proficiency, and 6) security respectively. Here we present the results of the linear regression model for these five independent variables as give by equation 1. The regression for Freelancer bidders and buyers were treated differently. Freelancer bidders are more akin to Mechanical Turk workers, as they bid on tasks posted by other market participants. Freelancer buyers, however, are more similar to those who post HITs on Mechanical Turk. While bidders represent supply, buyers create demand. Table 2 presents the results of the regression for all three dependent variables: 1) Mechanical Turk workers, 2) Freelancer bidders, and 3) Freelancer buyers. The cells in the table represent the p-value for which the specific indicator was significant in the best fit model⁶. NS indicates that the indicator was in the best fit model but was not statistically significant. More detailed information on the regression including estimates is available in Appendix C.

The first regression model had Mechanical Turk workers as the dependent variable. The adjusted R-square value was 0.9507, p-value=2.070e-08. The best fit for the model was given by digital economy ranking, connectivity and technology infrastructure, social and cultural environment, legal environment, government policy and vision, consumer and business adoption, English proficiency scores (reading, listening, speaking, and writing, number of Internet users, population, population percentage of women, voice and accountability, secure Internet servers, and secure Internet servers by population; adjusted R-square=0.9706, p-value<2.2e-16. (More details on the regression model are available in Appendix C, table 6.)

⁶Best fit model indicates the subset of indicators for which the corresponding linear regression obtained the highest adjusted R-square value.

Table 4: Exploratory Factor Analysis for Freelancer Bidder

Freelancer Bidder	Factor 1	Factor 2	Factor 3	Uniqueness
Connectivity & technology	0.926	0.119		0.126
Social & cultural environment	0.914	0.205		0.113
Government policy & vision	0.918	0.157		0.129
Consumer & business adoption	0.954	0.173		0.060
English reading	0.317	0.923		0.046
English writing	0.351	0.803		0.228
Number of Internet users			0.997	0.005
Population	-0.297	0.128	0.753	0.328
Export of ICT services	0.281	0.459	0.330	0.601
% export of ICT services		0.444		0.798
Rule of law	0.951	0.142		0.075
Government Effectiveness	0.947	0.166		0.075
Control of corruption	0.969			0.053
SIS	0.328		0.519	0.617
SIS by population	0.761	0.156		0.393
Proportion Variance	0.485	0.141	0.131	
Cumulative Variance	0.485	0.625	0.757	

The second run of the regression model had Freelancer bidders as the dependent variable. This regression gave an adjusted R-square value of 0.9306, p-value=5.035e-15. The best fit was given by connectivity and technology infrastructure, social and cultural environment, government policy and vision, English proficiency scores (reading, writing), number of Internet users, population, export of ICT services, percentage export of ICT services, rule of law, government effectiveness, control of corruption, secure Internet servers, and secure Internet servers by population; adjusted R-square=0.9481, p-value<2.2e-16. (More details on the regression model are available in Appendix C, table 7.)

The last run of the regression model had Freelancer buyers as the dependent variable. This model gave an adjusted R-square value of 0.9538, p-value<2.2e-16. The best fit was given by digital economy ranking, connectivity and technology infrastructure, business environment, social and cultural environment, government policy and vision, consumer and business adoption, TI corruption index, English proficiency scores (reading, listening), number of Internet users, population, population percentage of women, percentage of urban population, rule of law, regulatory quality, voice and accountability, control of corruption, secure Internet servers, secure Internet servers by population; adjusted R-square value=0.9625, p-value<2.2e-16. (More details on the regression model are available in Appendix C, table 8.)

Many of the indicators were highly correlated. Thus, the number of macro-level indicators can be condensed to smaller subset using exploratory factor analysis (EFA). We conducted a scree test to identify the optimum number of factors, Appendix B. Only the indicators that were present in the respective best fit models were considered. We considered the number of factors that give an eigenvalue greater than 1. Thus, there were four factors for Mechanical Turk and Freelancer Buyers, while they were three factors for Freelancer Bidders. The factor loadings for Mechanical

Table 5: Exploratory Factor Analysis for Freelancer Buyer

Freelancer Buyer	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
Digital economy ranking	0.961	0.197		0.184	0.005
Connectivity and technology	0.910	0.146		0.205	0.102
Business environment	0.916	0.115			0.141
Social & cultural environment	0.896	0.227		0.235	0.085
Government policy and vision	0.921	0.154		0.192	0.089
Consumer & business adoption	0.942	0.194		0.219	0.025
TI corruption index	0.946				0.100
English reading	0.312	0.907			0.078
English listening	0.414	0.818	-0.185	-0.100	0.115
Number of Internet users		-0.124	0.949		0.076
Population	-0.250		0.817	-0.204	0.228
Population % of women		0.451			0.794
% of urban population	0.554		-0.172	0.292	0.578
Rule of law	0.961	0.146		-0.110	0.043
Regulatory quality	0.914	0.199			0.117
Voice and accountability	0.747	0.458			0.221
Control of corruption	0.977				0.029
SIS	0.313		0.456	0.328	0.579
SIS by population	0.752	0.187		0.150	0.374
Proportion Variance	0.558	0.117	0.099	0.027	
Cumulative Variance	0.558	0.675	0.774	0.801	

Turk, Freelancer Bidders, and Freelancer Buyers are given in tables 3-5 respectively.

5 Discussion

In this paper, we empirically examine a macroeconomic model of crime in crowd-sourced labour markets. We investigate two research threads. First, we addressed the macro-level difference between two market equilibria, for crowd-sourced labor markets: 1) high enforcement low crime, 2) and low enforcement high crime. While Amazon’s Mechanical Turk service is an example of the former, Freelancer represents the latter. Secondly, we examined the macro-level differences between stakeholders in the Freelancer market, i.e. the those who bid on tasks, suppliers, and who created the tasks or demand.

We begin by identifying the macro-level indicators that encourage overall participation in crowd-sourced labor markets. We conducted an exploratory factor analysis to identify how the different indicators for the best fit models related to each other, tables 3-5. The factor loadings for Mechanical Turk as well as Freelancer were similar. While for Mechanical Turk and Freelancer buyers there were four factors, for Freelancer bidders there were three. The factor loadings on the fourth factors as well as the variance explained was low; 0.024 for Mechanical Turk and 0.027 for Freelancer Buyer. Thus, we assume that there are typically three factors that drive participation in either of these markets. The least amount of variance is explained by the third, and last, factor that constitutes population, number of Internet users, and SIS. This factor essentially indicates a cyber ready labor force. Thus, availability of a labor force with the essential skill set is required but not an adequate predictor of participation in crowd-sourced labor markets.

The second factor explains more variance and is characterized mostly by English language proficiency. English language proficiency is required, not only for completing a significant percentage of tasks, but even to understand the solicitations since most of the posts are made in English. The degree of English proficiency corresponds to the tasks afforded. A specific level of proficiency is not required for Freelancer bidding, ability to understand spoken English is required for Freelancer buying, and proficiency in listening as well as speaking is needed for Mechanical Turk.

The first factor explained most of the difference in the variance. It essentially constituted access to affordable Internet, both in terms of quantity and quality. Higher Internet penetration would allow a greater proportion of population to get online. However, Internet literacy is not facilitated by the ability to get online, but also the quality of the bandwidth available, e.g. broadband speed, security, reliability. Available bandwidth must be utilized, thus adoption both by individuals and businesses must be facilitated by public policies and private enterprise.

While there were broad factors that encouraged participation in crowd-sourced labor markets, there were differences on specific indicators that make the three big factors, shifting the market equilibria: high enforcement low crime market vs. low enforcement high crime markets. To examine these differences, we compare the best fit models of both markets, Mechanical Turk and Freelancer, bidders as well as buyers.

GDP per capita and GDP per capita by PPP were not in the best fit models of either Mechanical Turk or Freelancer, table 2. This finding is different from that observed for massive copyright infringement [40]. This indicates the importance of the direct financial rewards, or the degree to which they supplement or complement participants’ incomes is equally relevant for both markets.

Table 2 also shows that population, number of Internet users, and number of SIS are statistically significant indicators of participation in crowd-sourced markets. Participation, in either Mechanical Turk or Freelancer, was positively correlated with these three indicators. The three indicators

together form an indicator of a cyber ready labor force. All three load together as the third factor, table 3-5. Thus, while a good indicator of participation, they do not account much for the variance of participation in either market.

Connectivity and technology, social and cultural environment, government policy and vision, and English proficiency scores (reading) were other indicators that were common in the best fit models of Mechanical Turk and Freelancer, bidder as well as buyer. These factors were not statistically significant for all three, however, these factors indicate a necessary, if not sufficient, indicators for increasing participation. These factors account for the most of the variance in the best fit models for both Mechanical Turk and Freelancer, bidders as well as buyers. Connectivity and technology, social and cultural environment, and government policy and vision loaded on the first factor while English proficiency loaded on the second, as seen by the results of EFA in tables 3-5.

There were no statistically significant indicators that were common between Mechanical Turk workers and Freelancer bidders. This finding is counterintuitive, as Mechanical Turk workers as well as Freelancer bidders are both suppliers in respective markets, responding to a demand created by buyers. The expectation would be that since the services provided are similar, the macro economic backgrounds would be too. However, for Mechanical Turk workers the overall state of the digital economy was important, mostly driven by Internet penetration, via broadband as well as mobile. Affordable access to high bandwidth that is reliable and secure is important. Freelancer bidders may not have lower quality or quantity of access. Participation as a Freelancer bidder is driven by social and cultural environment, having lower technical expertise and English proficiency than their Mechanical Turk counterparts. Social and cultural environment extends to legal framework and the ability of the government to implement it. It is important to note that this is different from perceptions of corruption, rather this is more an indicator of general perceptions of rules being followed in society and penalties being imposed for deviations.

Statistically significant common indicators between Freelancer bidders and buyers are also limited. The one common indicator was social and cultural environment. Both bidding and buying, then requires a minimum level of technical expertise and general education. However, the switch from bidder to buyer demands several other resources. Participation as buyers requires access to affordable bandwidth. Solicitation of jobs would probably be a part of the larger business strategy. Thus, business environment is also important and so is consumer and business adoption. English language proficiency is again more important for buyers than for bidders. Percentage of urban population is also an indicator of buying. Urban areas are likely to have higher quality and quantity of Internet penetration, so this is not surprising.

However, Mechanical Turk participation, which also requires access to affordable to bandwidth, is not dependent on urban population. This difference becomes more clear when we consider that buying is also a function of population percentage of women and corruption. These indicators together suggest that buying in Freelancer is explained by social disorganization theory in criminology [47]. Social disorganization theory considers crime to be a function of the breakdown of community structures, as measured by corruption, heterogeneity of population and urbanization.

This is different from participation as bidder, which is dependent on rule of law and the effectiveness of the government to enforce the rules. Thus, for bidders deterrence based efforts such as increasing penalties on increased enforcement might be successful at decreasing participation [20]. However, the long term impact of such measures would be limited [43]. Without the demand for Freelancer services being alleviated, there is likely to be a displacement effect, where bidders from countries with a lax legal framework would increasingly participate [44].

6 Conclusion and Future Work

In this work we examine the macro-level factors that encourage participation in crowd-sourced labor markets online. We differentiate these crowd-sourced labor markets as those with lower number of illegitimate tasks vs. those where a majority of tasks support criminal enterprise online. In previous work we posited that ICT markets would tend to exist in one of two equilibria (i.e, high enforcement and low crime or low enforcement and high crime). We considered the specific instance of Amazon’s Mechanical Turk representing high enforcement low crime and Freelancer representing low enforcement high crime. We identified the macro-level factors that appear to facilitate either of these two market equilibria.

A complementary approach would be to identify high enforcement/low crime and low enforcement/high crime countries, and then examine the proportion of participation in Mechanical Turk by those countries as opposed to in Freelancer. This would be addressed in future research. Adoption of Mechanical Turk/Freelancer may also be driven by the dynamics of social networks. A macro-level investigation examines variables that facilitate such dynamics. This, however, does not directly address why certain networks may become more popular in specific countries. To the degree that adoption is driven by trust and just plain awareness, physical social networks may in fact influence the decision to participate in a specific crowd-sourced labor market. However, given that crowd-sourced labor markets are not like typical social networks, such as Facebook or telephone networks, we assume that the impact of network effects is trivial. We also do not address the intentionality of participants from either a rational microeconomic or a boundedly rational behavioral perspective. Individual motivations can be pursued through a survey based study, and should be addressed by future research.

Participation in crowd-sourced labor markets requires three factors: 1) affordable access to reliable and secure Internet, 2) English language proficiency, and 3) availability of a cyber ready labor force. A high enforcement low crime equilibrium has several additional characteristics. The digital economy of the participants’ local jurisdictions must be thriving so as to provide access to affordable, reliable, and secure bandwidth. Participants must have high technical as well as language skills. Adoption of ICTs by individuals as well as businesses in the local jurisdiction must be sufficient and facilitated by government policies.

There were several statistically significant indicators that were common between Mechanical Turk and Freelancer buyers: digital economy ranking, connectivity and technology, consumer and business adoption, and English proficiency scores (listening). So why do more Mechanical Turk workers not gravitate towards Freelancer? A possible explanation is that Mechanical Turk workers have higher level of English proficiency as well as formal education and Internet literacy. Simultaneously, Freelancer buyers come from jurisdictions with higher urban population, more corruption, as well as those conducive to a business enterprise that is not necessarily ethical.

A low enforcement high crime equilibrium is facilitated by a poor legal framework. Freelancer bidders’ local jurisdictions often do not have an effective or comprehensive legal framework. Even when the legal framework is adequate, prosecution is denied due to corruption in the local jurisdictions of Freelancer buyers. Participation through bidding in these markets can be alleviated, in the short term, through deterrence based strategies, such as increasing penalties or higher enforcement. Yet the very factors of corruption and lack of consistent rule of law limits the efficacy of these short-term efforts. Long terms success of deterrence would be limited. Thus, short terms policies must be complemented by long-term strategies. Two strategies are addressing participation through buying (i.e. by alleviating the demand for criminal goods), and changing the underlying macro-level

factors. The second requires shifting the underlying macro-level factors of underdeveloped ICT markets, inadequate language skills (in English) as well as Internet literacy.

Acknowledgements

We would like to thank Prof. Panagiotis G. Ipeirotis who made the demographic data publicly available. We also thank Prof. Stefan Savage's research group at UCSD for providing us the data on Freelancer. Finally, we thank the Stat/Math center at Indiana University for their insight on the statistical analysis. Any mistakes in this paper are the authors' own responsibility.

This presentation of this research was made possible by funding from the Volkswagen Foundation. This material is based upon work supported by the National Science Foundation under award number 0916993. Any opinions, findings, and conclusions or recommendations expressed in this presentation are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

- [1] I. Al-Jabri and A. Abdul-Gader. Software copyright infringements: an exploratory study of the effects of individual and peer beliefs. *Omega*, 25(3):335–344, 1997.
- [2] R. Anderson. Why information security is hard-an economic perspective. In *Computer Security Applications Conference, 2001. ACSAC 2001. Proceedings 17th Annual*, pages 358–365. IEEE, 2001.
- [3] A. Bernard and M. Busse. Who wins the olympic games: Economic resources and medal totals. *Review of Economics and Statistics*, 86(1):413–417, 2004.
- [4] J. Bhagwati and B. Hansen. A theoretical analysis of smuggling. *The Quarterly Journal of Economics*, pages 172–187, 1973.
- [5] R. Bursik Jr and H. Grasmick. Economic deprivation and neighborhood crime rates, 1960-1980. *Law and Society Review*, pages 263–283, 1993.
- [6] E. Bursztein, S. Bethard, C. Fabry, J. Mitchell, and D. Jurafsky. How good are humans at solving captchas? a large scale evaluation. In *Security and Privacy (SP), 2010 IEEE Symposium on*, pages 399–413. IEEE, 2010.
- [7] M. Chamlin and J. Cochran. Assessing messner and rosenfeld's institutional anomie theory: A partial test*. *Criminology*, 33(3):411–429, 1995.
- [8] M. Chamlin and J. Cochran. Social altruism and crime*. *Criminology*, 35(2):203–226, 1997.
- [9] Y. Chen and I. Png. Software pricing and copyright enforcement: private profit vis-a-vis social welfare. In *Proceedings of the 20th international conference on Information Systems*, pages 119–123. Association for Information Systems, 1999.
- [10] K. Choo and R. Smith. Criminal exploitation of online systems by organised crime groups. *Asian journal of criminology*, 3(1):37–59, 2008.

- [11] N. Christin, S. Egelman, T. Vidas, and J. Grossklags. Its all about the benjamins: An empirical study on incentivizing users to ignore security advice. *Financial Cryptography and Data Security*, pages 16–30, 2012.
- [12] L. Cohen and M. Felson. Social change and crime rate trends: A routine activity approach. *American sociological review*, pages 588–608, 1979.
- [13] M. Colvin, F. Cullen, and T. VEN. Coercion, social support, and crime: An emerging theoretical consensus*. *Criminology*, 40(1):19–42, 2002.
- [14] W. Easterly and M. Sewadeh. *Global development network growth database*. World Bank Group, 2001.
- [15] L. Edlund, H. Li, J. Yi, and J. Zhang. Sex ratios and crime: Evidence from chinas one-child policy. *IZA DP*, (3214), 2007.
- [16] M. Fortunato. Let’s not go crazy: Why lenz vs. universal music corp. undermines the notice and takedown process of the digital millennium copyright act. *Journal of Intellectual Property Law*, 17:147–445, 2009.
- [17] J. Franklin, V. Paxson, A. Perrig, and S. Savage. An inquiry into the nature and causes of the wealth of internet miscreants. In *ACM Conference on Computer and Communications Security (CCS)*, 2007.
- [18] E. Gal-Or and A. Ghose. The economic consequences of sharing security information. *Economics of information security*, pages 95–104, 2004.
- [19] V. Garg, N. Husted, and J. Camp. Smuggling theory approach to organized digital crime. In *eCrime Researcher’s Summit*. IEEE, 2011.
- [20] G. Higgins, A. Wilson, and B. Fell. An application of deterrence theory to software piracy. *Journal of Criminal Justice and Popular Culture*, 12(3):166–184, 2005.
- [21] T. Holz, M. Engelberth, and F. Freiling. Learning more about the underground economy: A case-study of keyloggers and dropzones. *Computer Security–ESORICS 2009*, pages 1–18, 2009.
- [22] J. Horton and L. Chilton. The labor economics of paid crowdsourcing. In *Proceedings of the 11th ACM conference on Electronic commerce*, pages 209–218. ACM, 2010.
- [23] P. Ipeirotis. Demographics of mechanical turk. *Center for Digital Economy Research, NYU Stern School of Business, Working paper*, 2010.
- [24] C. Kanich, S. Checkoway, and K. Mowery. Putting out a hit: crowdsourcing malware installs. In *Proceedings of the 5th USENIX Workshop on Offensive Technologies*, 2011.
- [25] M. Katz and C. Shapiro. Technology adoption in the presence of network externalities. *The journal of political economy*, pages 822–841, 1986.
- [26] D. Kaufmann, A. Kraay, and M. Mastruzzi. The Worldwide Governance Indicators: Methodology and Analytical Issues. *SSRN eLibrary*, 2010.

- [27] S. Krishnamurthy and A. Tripathi. Bounty programs in free/libre/open source software. *The Economics of Open Source Software Development*. Hrsg.: Schröder, P., Elsevier, Amsterdam, S, pages 165–184, 2006.
- [28] J. Kwon and M. Johnson. An organizational learning perspective on proactive vs. reactive investment in information security. In *Workshop on Economics of Information Security*, 2011.
- [29] M. Lelarge. Economics of malware: Epidemic risks model, network externalities and incentives. In *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, pages 1353–1360. IEEE, 2009.
- [30] Z. Li, Q. Liao, and A. Striegel. Botnet economics: uncertainty matters. *Managing Information Risk and the Economics of Security*, pages 245–267, 2009.
- [31] W. Mason and S. Suri. Conducting behavioral research on amazons mechanical turk. *Behavior Research Methods*, pages 1–23, 2010.
- [32] W. Mason and D. Watts. Financial incentives and the performance of crowds. *ACM SIGKDD Explorations Newsletter*, 11(2):100–108, 2010.
- [33] S. Messner and R. Rosenfeld. Crime and the american dream. No.: ISBN 0-534-20106-7, page 143, 1994.
- [34] C. Miller. The legitimate vulnerability market: Inside the secretive world of 0-day exploit sales. In *In Sixth Workshop on the Economics of Information Security*. Citeseer, 2007.
- [35] T. Moore and R. Clayton. An empirical analysis of the current state of phishing attack and defence. In *Workshop on the Economics of Information Security*. Citeseer, 2007.
- [36] T. Moore, R. Clayton, and R. Anderson. The economics of online crime. *The Journal of Economic Perspectives*, 23(3):3–20, 2009.
- [37] M. Motoyama, K. Levchenko, C. Kanich, D. McCoy, G. Voelker, and S. Savage. Re: Captchas—understanding captcha-solving services in an economic context. In *USENIX Security Symposium*, volume 10, 2010.
- [38] M. Motoyama, S. Savage, and G. M. Voelker. Dirty jobs : The role of freelance labor in web service abuse. In *Proceedings of Usenix Security Symposium*, 2011.
- [39] I. Nagel and J. Hagan. Gender and crime: Offense patterns and criminal court sanctions. *Crime & Just.*, 4:91, 1983.
- [40] C. Osorio. A contribution to the understanding of illegal copying of software: Empirical and analytical evidence against conventional wisdom. In *Program on Internet and Telecoms Convergence*. MIT, 2002.
- [41] A. Ozment. Bug auctions: Vulnerability markets reconsidered. In *Third Workshop on the Economics of Information Security*, 2004.
- [42] M. Pitt. Smuggling and price disparity. *Journal of International Economics*, 11(4):447–458, 1981.

- [43] I. Png and C. Wang. The deterrent effect of enforcement against computer hackers: Cross-country evidence. In *WEIS 2007-Sixth Workshop on Economics of Information Security*, 2007.
- [44] I. Png, C. Wang, and Q. Wang. The deterrent and displacement effects of information security enforcement: International evidence. *Journal of Management Information Systems*, 25(2):125–144, 2008.
- [45] T. Pratt and F. Cullen. Assessing macro-level predictors and theories of crime: A meta-analysis. *Crime and Justice*, pages 373–450, 2005.
- [46] J. Ross, L. Irani, M. Silberman, A. Zaldivar, and B. Tomlinson. Who are the crowdworkers?: shifting demographics in mechanical turk. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pages 2863–2872. ACM, 2010.
- [47] R. Sampson and W. Groves. Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, pages 774–802, 1989.
- [48] J. Sanchez. Sopa, internet regulation and the economics of piracy. Technical report, CATO Institute, 2012.
- [49] J. Savolainen. Inequality, welfare state, and homicide: Further support for the institutional anomie theory*. *Criminology*, 38(4):1021–1042, 2000.
- [50] A. Shostack and A. Stewart. *The new school of information security*. Addison-Wesley Professional, 2008.
- [51] B. Stone-Gross, R. Abman, R. Kemmerer, C. Kruegel, D. Steigerwald, and G. Vigna. The underground economy of fake antivirus software. In *Workshop on the Economics of Information Security*, 2011.
- [52] R. Thomas and J. Martin. The underground economy: priceless. *USENIX; login*, 31(6):7–16, 2006.
- [53] E. I. Unit. Digital economy rankings 2010 beyond e-readiness. Technical report, Economist Intelligence Unit and The IBM Institute for Business Value, 2010.
- [54] H. Varian. Economics of information technology. *University of California, Berkeley*, 2001.
- [55] H. Varian. System reliability and free riding. *Economics of Information Security*, pages 1–15, 2004.
- [56] V. Venkatesh and M. Morris. Why don't men ever stop to ask for directions? gender, social influence, and their role in technology acceptance and usage behavior. *Mis Quarterly*, pages 115–139, 2000.
- [57] L. von Ahn. Human computation. In *Design Automation Conference, 2009. DAC'09. 46th ACM/IEEE*, pages 418–419. IEEE, 2009.
- [58] L. von Ahn, M. Blum, N. Hopper, and J. Langford. Captcha: Using hard ai problems for security. In *Proceedings of the 22nd international conference on Theory and applications of cryptographic techniques*, pages 294–311. Springer-Verlag, 2003.

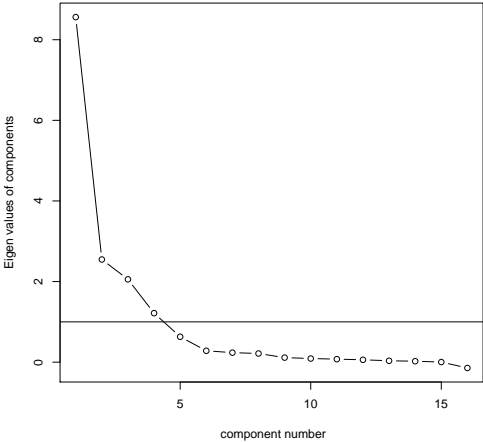
- [59] I. Walker and T. Pettigrew. Relative deprivation theory: An overview and conceptual critique. *British Journal of Social Psychology*, 23(4):301–310, 1984.
- [60] M. Yar. The novelty of ‘cybercrime’. *European Journal of Criminology*, 2(4):407, 2005.
- [61] J. Zhuge, T. Holz, C. Song, J. Guo, X. Han, and W. Zou. Studying malicious websites and the underground economy on the chinese web. *Managing Information Risk and the Economics of Security*, pages 225–244, 2009.

A List of Countries

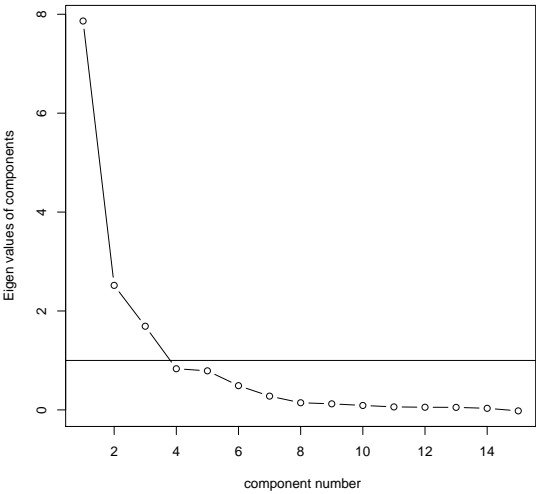
The following countries were considered in our analysis: United States, United Kingdom, Canada, India, Bangladesh, Philippines, Pakistan, Romania, Kenya, Nigeria, Sri Lanka, Indonesia, Australia, Ukraine, Malaysia, Egypt, South Africa, China, Bulgaria, Vietnam, Nepal, Thailand, Singapore, Argentina, United Arab Emirates, Spain, Serbia, Italy, Ghana, Israel, Mauritius, Russian Federation, Germany, New Zealand, Sweden, France, Jamaica, Turkey, Ireland, Portugal, Mexico, Brazil, Uganda, Greece, Uruguay, Saudi Arabia, Macedonia, Croatia, Trinidad And Tobago, Poland, Netherlands, Oman, Hong Kong, Morocco, Colombia, Jordan, Cameroon, Tunisia, Estonia, Peru, Japan, Moldova, Hungary, Costa Rica, Cyprus, Switzerland, Bosnia And Herzegovina, Honduras, Belarus, Czech Republic, Kuwait, Tanzania, Belgium, Norway, Puerto Rico, Bahrain, Iran, Denmark, Zimbabwe, Senegal, Algeria, Lebanon, Chile, Venezuela, Dominican Republic, Austria, Latvia, Finland, Lithuania, Madagascar, Qatar, Malta, Botswana, Slovenia, Armenia, Georgia, Bolivia, Panama, Slovak Republic, Taiwan, Korea, Seychelles, Iceland, Belize, Gibraltar, Cambodia, Luxembourg, and Albania.

B Scree Test

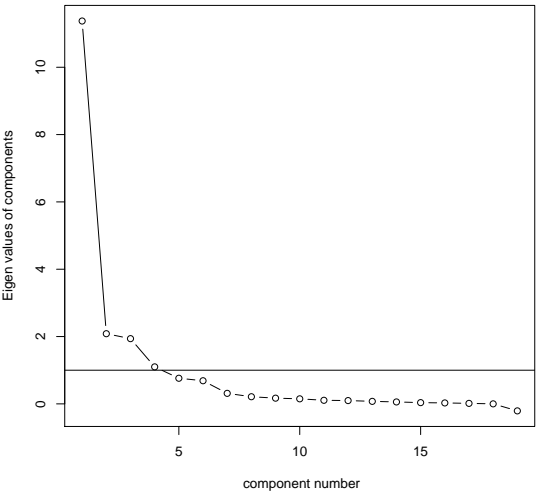
Mechanical Turk



Freelancer Bidder



Freelancer Buyer



C Best Fit Regression Models

Table 6: Mechanical Turk

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.571e+01	5.977e+01	0.597	0.55541
Digital economy rankings	5.578e+01	2.491e+01	2.239	0.03393 *
Connectivity & technology	-1.826e+01	6.405e+00	-2.851	0.00843 **
Social & cultural environment	1.185e+01	7.431e+00	1.595	0.12276
Legal environment	9.148e+00	5.706e+00	1.603	0.12094
Government policy & vision	-1.358e+01	7.302e+00	-1.860	0.07419 .
Consumer & business adoption	-2.499e+01	9.189e+00	-2.720	0.01148 *
English reading	-7.021e+00	3.631e+00	-1.934	0.06413 .
English listening	1.662e+01	4.542e+00	3.660	0.00113 **
English speaking	-7.782e+00	3.527e+00	-2.207	0.03637 *
English writing	-6.019e+00	4.082e+00	-1.475	0.15229
Number of Internet users	-1.256e-06	8.949e-08	-14.037	1.21e-13 ***
Population	3.469e-07	2.142e-08	16.196	4.22e-15 ***
Population % of women	-1.302e+00	1.068e+00	-1.219	0.23378
Voice & accountability	-4.775e-01	2.464e-01	-1.938	0.06361 .
SIS	1.518e-03	5.284e-05	28.730	< 2e-16 ***
SIS by population	-2.213e-02	7.626e-03	-2.902	0.00745 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.84 on 26 degrees of freedom

Multiple R-squared: 0.9818, Adjusted R-squared: 0.9706

F-statistic: 87.72 on 16 and 26 DF, p-value: < 2.2e-16

Table 7: Freelancer Bidder

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.230e+04	5.198e+04	0.814	0.42028
Connectivity & technology	5.361e+03	5.100e+03	1.051	0.29909
Social & cultural environment	-2.333e+04	8.337e+03	-2.799	0.00765 **
Government policy & vision	7.139e+03	6.209e+03	1.150	0.25655
Consumer & business adoption	1.494e+04	9.029e+03	1.654	0.10532
English reading	5.980e+03	3.437e+03	1.740	0.08903 .
English writing	-5.294e+03	3.905e+03	-1.356	0.18224
Number of Internet users	-2.130e-03	1.478e-04	-14.411	< 2e-16 ***
Population	6.555e-04	3.581e-05	18.306	< 2e-16 ***
Export of ICT services	-6.188e-06	5.831e-06	-1.061	0.29454
% export of ICT services	1.305e+03	6.683e+02	1.953	0.05737 .
Rule of law	1.463e+03	6.132e+02	2.386	0.02151 *
Government effectiveness	-1.413e+03	5.714e+02	-2.472	0.01746 *
Control of corruption	-8.318e+02	5.945e+02	-1.399	0.16890
SIS	1.293e+00	7.823e-02	16.529	< 2e-16 ***
SIS by population	-1.439e+01	9.240e+00	-1.558	0.12666

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22760 on 43 degrees of freedom

Multiple R-squared: 0.9615, Adjusted R-squared: 0.9481

F-statistic: 71.62 on 15 and 43 DF, p-value: < 2.2e-16

Table 8: Freelancer Buyer

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.009e+03	5.832e+03	0.516	0.608966
Digital Economy Ranking	9.168e+03	2.941e+03	3.117	0.003468 **
Connectivity & technology	-2.651e+03	6.675e+02	-3.972	0.000307 ***
Business environment	-2.179e+03	7.743e+02	-2.815	0.007697 **
Social & cultural environment	-1.417e+03	6.419e+02	-2.207	0.033410 *
Government policy & vision	-1.534e+03	6.760e+02	-2.270	0.028994 *
Consumer & business adoption	-2.490e+03	9.311e+02	-2.674	0.010990 *
TI corruption index	7.181e+01	4.079e+01	1.760	0.086406 .
English reading	-3.905e+02	2.482e+02	-1.573	0.123995
English listening	7.278e+02	2.329e+02	3.125	0.003396 **
Number of Internet users	-5.233e-05	8.128e-06	-6.438	1.43e-07 ***
Population	1.569e-05	2.015e-06	7.790	2.15e-09 ***
Population % of women	-2.250e+02	9.972e+01	-2.256	0.029926 *
% of urban population	4.324e+01	1.725e+01	2.506	0.016598 *
Rule of law	4.521e+01	4.000e+01	1.130	0.265553
Regulatory Quality	7.577e+01	3.984e+01	1.902	0.064781 .
Voice & accountability	-2.351e+01	1.849e+01	-1.271	0.211280
Control of corruption	-1.123e+02	5.304e+01	-2.117	0.040832 *
SIS	1.416e-01	5.083e-03	27.864	< 2e-16 ***
SIS by population	-1.183e+00	7.489e-01	-1.579	0.122609

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1434 on 38 degrees of freedom

Multiple R-squared: 0.975, Adjusted R-squared: 0.9625

F-statistic: 78.02 on 19 and 38 DF, p-value: < 2.2e-16