RAND EUROPE

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: Jump to Page 1 ▼

Support RAND

Browse Reports & Bookstore

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore RAND Europe

View document details

R E P O R T

RAND EUROPE

Measuring IPR infringements in the internal market

Development of a new approach to estimating the impact of infringements on sales

Stijn Hoorens, Priscillia Hunt, Alessandro Malchiodi,
Rosalie Liccardo Pacula, Srikanth Kadiyala,
Lila Rabinovich, Barrie Irving

RAND EUROPE

Preface

There is no shortage of estimates of the scale and impacts of IPR infringements. However, there is little consensus on the accuracy or reliability of these numbers. In absence of a robust evidence base, it is difficult to debate the effectiveness of government efforts to regulate intellectual property rights (IPRs) or address the impacts of infringements. Therefore, the European Commission aims to develop and implement a system that monitors trends in this area.

The study was commissioned by the Internal Market and Services DG of the European Commission. It was set up to assist the Commission in the development of a methodology that would quantify the scope, scale and impact of IPR infringements on the European economy. In this report we offer the 'building blocks' for such a methodology that strives to be consistent, robust, feasible and reliable in measuring the scale of this phenomenon.

Based on an extensive review of the literature, we propose a methodology for measuring trends of the lost revenues due to IPR infringements in markets of counterfeited products. While the methodology presents a promising approach to the problem, a number challenges remain and it needs further testing. The report offers various recommendations for next steps to take this approach to the next level.

The study has been a joint effort of RAND researchers based in Brussels, Cambridge (UK), Washington DC and Santa Monica (US). This report has been peer-reviewed in accordance with RAND's quality assurance standards. The document should be relevant to policymakers with an interest in measuring the extent and scale of IPR infringements at national, European or global level. Moreover, the report will be useful to industry representatives or analysts interested in the impact of this phenomenon at firm level.

RAND Europe is an independent not-for-profit policy research organisation that aims to improve policy and decision making in the public interest, through research and analysis. RAND Europe's clients include European governments, institutions, NGOs and firms with a need for rigorous, independent, multidisciplinary analysis. For more information about RAND or this document, please contact:

Stijn Hoorens
RAND Europe
Rue de la Loi 82
1040 Brussels, Belgium
e-mail: hoorens@rand.org
tel. +32 2966 2400

Contents

Executive summary

Counterfeiting is not a new phenomenon. For centuries, artists or inventors have seen their creations and products copied without their permission. However, it is with the trends of globalisation, the integration of markets and the rise of the Internet economy in recent decades that violations of Intellectual Property Rights (IPRs) have become more widespread. Easy access to computers, Internet and other technological developments facilitate duplication of designs, labels, logos, packaging and documentation with speed, accuracy and relative anonymity.

IPRs refer to protections granted to firms and/or individuals who are the creators of ideas, products, or methods that allow the creators/inventors a period of time in which they can earn exclusive returns on these intangible and tangible products as a way of rewarding them for the risky investment they initially made. Counterfeiting these products or sharing creative content without permission of the creators infringes upon these intellectual property rights. We distinguish two types of infringements: counterfeiting of physical products and unauthorised use of protected content (UUPC), which is commonly referred to as piracy.

It is argued that 'victims' of counterfeiting and UUPC could face considerable economic, health and safety impacts. Many of these will impact the IPR holders, for example if consumers purchase these counterfeited or pirated substitutes instead of the legitimate products. In turn for consumers, their health or safety may be compromised. Some argue on the other hand that some forms of IPR infringements may even have positive externalities.

There is no shortage of estimates of the extent of IPR infringements, and there is some empirical evidence of negative impacts of these infringements in specific sectors. However, most of these efforts lack a transparent methodology, suffer from serious methodological or data limitations or are funded by stakeholders in the debate. This means that the resulting estimates must be heavily caveated and qualified, putting into question the extent to which they are useful to governments and firms trying to understand and tackle the phenomenon. Without objective and reliable estimates of the extent of IPR violations it is difficult to debate these claims.

Given the intensity of these debates, an objective and evidence-based approach towards measuring the scale and impact of the phenomenon has become more important than ever. This study was set up to assist the European Commission in the development of a methodology that would quantify the scope, scale and impact of IPR infringements on the European economy in the Internal Market. This study is the first stage in an attempt to continuously assess the problem and to develop evidence-based policies in the area of intellectual property rights. In this report we offer the "building blocks" for a methodology that is consistent, robust, feasible and reliable in measuring the size of counterfeiting and UUPC. Further testing of the methodology is recommended in multiple industry sectors to better understand the scope and scale of the

problem before it is possible to move to the next stage, which would involve assessing the impact of counterfeiting and UUPC on industries, government and public health or safety.

In this report we aim to address a number of research questions that help to achieve this goal.

What can we learn from previous efforts about the drivers and impacts of IPR infringements?

In order to develop a theoretical basis for a method to estimate the extent of IPR infringements, it is important to understand the factors that encourage suppliers to offer products that are in violation of these rights or drive consumers to buy them. Some of these drivers of supply and demand of counterfeit products or UUPC are summarised in the table below.

Macro-level drivers of supply	Macro-level drivers of demand
• The growing prevalence of digital and networked technologies	• Social acceptance to buy products that violate intellectual property rights
• The globalization of trade, the growing importance of international brands	• Limited availability of authentic goods
• The presence of large integrated markets supporting free trade	• The high price of authentic goods
• Low or weak enforcement of penalties targeting violators of IPR infringements	• The rising quality of counterfeit goods.
• The growing presence and involvement of organized crime in the production and distribution of counterfeited and pirated goods	
• Industry-specific factors.	

The production or consumption of products that infringe IPRs could have important implications for rights' holders, consumers, governments, employees, etc. There have been a number of attempts to estimate the variety of impacts, using different methodologies. They range from an annual $ 77.5bn in lost tax revenues in the G20 to 120,085 jobs lost in the motion picture industry. By offering a sample of these estimates, this report provides an indication of the variation in attempts and some indicative empirical evidence for the breadth of the effects of IPR infringements and their order of magnitude.

The validity and reliability of these estimates have been extensively challenged in previous studies. Either they tend to lack the necessary transparency or, when rigorously describing their methodologies, they have been criticized for some of their assumptions. Given the poor understanding of the extent of IPR infringements in the internal market, which is a necessary basis for estimating their effects, we focus in this study on developing an approach to measure trends in counterfeit and UUPC markets.

What are the strengths and weaknesses of existing methodologies that have been applied to measure the scope, size and impact of IPR infringements?

We identified nearly 250 publications addressing the issues relevant for this study, 80 of which were analysed in detail. We focused on studies that provided original attempts to quantitatively estimate the scope, size or effects of counterfeiting and/or UUPC, in any geographical area, and for one or more products. For each source we assessed the robustness and suitability of the methodological approach for the purpose of our study. We have drawn a number of lessons from this review:

- **Proxy indicators are needed to assess the magnitude of illicit markets.** In illicit markets, it is not possible to observe either demand or supply of counterfeits and UUPC directly. This means that proxy indicators, i.e. indirect measures that approximate or represent the real phenomena will be required to estimate their production, trade and consumption. Moreover, it will be necessary to triangulate information from alternative approaches and sources. Not all proxy indicators are equally suitable to provide reliable estimates.

- **Data sources are primarily based on consumer surveys and seizures.** Most data sources for estimating the size of these markets draw largely on self-reported information from consumers and suppliers as well as law enforcement data on known seizures and law suits. Some efforts use these data sources in conjunction with information obtained through more sophisticated, but resource-intensive approaches such as mystery shopping or sting operations.

- **There is little convergence on methodology in counterfeiting literature.** While numerous attempts to approximate the scale of counterfeiting have been made, there has been little convergence on a preferred methodology that can be broadly applied across industry sectors: innovation in methods and forms of collecting data continue to evolve. Furthermore it is often difficult to assess the quality of specific studies, as there is little transparency in assumptions and data values or sources; often for good reasons.

- **There is more convergence in the literature on copyright infringements.** Most approaches focus on "online piracy" these days. The absence of tangible goods has fundamental consequences for the distribution channels of UUPC. Estimates for these copyright infringements typically utilise survey methods and download or Internet traffic statistics. However, there is lack of clarity and consistency in how extrapolations to specific markets or countries are performed, in large part because the literature has not yet reached consensus on what drives copyright infringements.

- **Much work is needed around consumers' substitution rates.** There is still considerable uncertainty about the extent to which consumers substitute legitimate products and for those that violate IPR. Assumptions on the substitution rate depend on the consumers' knowledge and assessment of the quality. These often vary considerably by product and/or they remain unknown.

- **Methods for extrapolating to other markets or countries lack clarity.** More serious consideration regarding how to aggregate findings for specific products across countries needs to be considered in light of the nuances of the different product markets. As it is likely to be unfeasible to collect empirical data on all products and all geographical areas, some aggregation will be required. However, regional and market-specificities may make extrapolation based on general assumptions across countries and product types not reasonable or reliable.

- **Reliability and rigour may need priority over coverage.** There seems to be scope for sacrificing coverage of products targeted by counterfeiters for reliability and in terms of developing a model that works for specific products, at least to start with. This is because the hidden nature of these markets requires that significant effort and learning is required when trying to measure these phenomena in a meaningful way across countries and product groups.

- **The literature reveals a preference for "market-based" approaches.** Many studies focus on lost sales to legitimate IPR holders. This can be considered as a proxy of the market size for

counterfeiting and UUPC, but also represents a first-order effect. Clarifying the distinction between size and effects from a conceptual perspective is not a crucial question for future efforts in this field. However, from a practical standpoint our review suggests that lost sales, measured in terms of quantities or revenues, are a sensible outcome to consider when estimating the size of counterfeiting and piracy using a "market-based" approach.

- **A first-cut quantitative analysis of impacts may not require complicated methodologies.** Given the current state of knowledge, studying the impact of counterfeiting and UUPC does not require sophisticated econometric techniques. The linear regressions found in the literature so far are a good starting point provided that the right interpretation is attached to them. The biggest challenge remains in obtaining reliable measures of the magnitude of counterfeiting and UUPC.

What does this mean for the development of a methodology to be applied by the European Commission to estimate the scale of IPR infringements in the internal market?

We conclude that there is no reliable or accepted method for estimating the size of counterfeiting and UUPC that is feasible for the purposes of producing annual measures for all the affected products or markets and in all countries. While different approaches, such as consumer surveys or mystery shopping, can provide useful insights towards understanding specific aspects of these markets, it appears there is no one-size-fits-all solution available.

A market-based approach to estimate lost sales to rights' holders seems a sensible approach to proxy the market for IPR infringements in the EU and as a first-order estimate of the effects. While lost sales or revenues only partially represent the potential impacts of infringements, it is a first step in developing a monitoring system for the measurement of other impacts more broadly (e.g. on innovation, growth and competitiveness, creativity and culture, public health and safety, employment, environment, tax revenues and crime).

What would be the characteristics and data requirements of such a methodology?

We propose and test a new methodology for estimating lost sales motivated by economic theory that has been applied to specific industries in a few instances. We propose to use firm forecasts combined with information in the literature on country level measures related to counterfeiting to understand counterfeiting trends. Our key insight is that the relationship between these country level measures and unexpected differences between firm forecasts and sales provides us a proxy for estimating trends in IPR violations.

The methodology we develop to estimate the size of the market for counterfeit goods is a supply-side approach making use of economic theories of monopolistic competition and differentiable goods. Counterfeiters are more attracted to markets where firms are able to extract some monopoly rents. This can either be due to product differentiation or because it is a true oligopoly. Our approach attempts to exploit this feature in its empirical strategy to estimate the size of the market.

The idea is to estimate from legitimate producers the post-hoc amount of "unmet demand" that they experience and use that as a proxy for total counterfeit products sold. We presume that rights holders who are targeted by counterfeiters are able to calculate this amount as the residual of their forecasted demand for their products net actual units sold.

If a leading firm in an oligopoly market is unable to sell the predicted quantity it projects, it is typically due to an unexpected shock that is observable after the fact, such as a shock in the

supply chain, poorly received advertising campaign, or even a financial crisis or earthquake that impacts the overall economy. Once these factors are taken into account, the revised predicted quantities look a lot more like the volumes actually sold. However, sometimes the revised projection still cannot account for the difference between revised forecasted sale and actual sales.

This unexplained unfulfilled demand, our model presumes, is due at least in part to IPR infringements. A statistical model then attempts to identify the portion of unexplained unfulfilled demand that is highly correlated with factors that drive IPR infringements of a particular product in a particular country. These factors may include: the rule of law, control of corruption, level of tourism, access to broadband Internet or government effectiveness. This approach implies a two-stage regression based on the steps outlined in the box below.

A two-staged approach to estimate sales lost due to IPR infringement

1) **First stage**: identify unexplained error
    a. Based on forecasts of quantities of specific products that firms expected to sell in a given time period, calculate the difference between the forecast and real quantities sold;
    b. Identify any "observable" reasons for "error";
    c. Remaining difference is the unexplained forecast error.
2) **Second stage**: estimate proportion of unexplained error that is caused by IPR infringements
    a. Collect indicators on observable factors for a specific year that have been identified in the literature as related to consumption and supply of products that infringe IPRs. These may include legal (e.g. rule of law), economic (e.g. international tourism) or technological (e.g. broad Internet access) factors;
    b. Correlate unexplained error to those second stage regression factors;
    c. Generate an estimate of the amount of unexplained forecasting error that can be predicted by supply and demand factors of IPR infringements.

The approach requires firm forecast data on products that are subject to IPR infringements, as well as the actual items sold in different countries. The first stage regression requires retrospective information on observable product-, firm- and/or market-specific factors that explain the error. These may include data on competitors' sales, GDP growth, consumer trust, foreign exchange rates, etc. The second stage regression requires annual descriptive statistics on factors related to IPR infringements in specific countries.

What does application of this methodology teach us about the scale and impacts of IPR infringements in the internal market?

This new methodology has been empirically tested using confidential data made available to us by a multinational technology firm producing consumer goods targeted by counterfeiters. Preliminary evidence suggests that the proposed alternative two-step methodology may be a fruitful avenue forward for monitoring trends in the overall size of counterfeit markets, particularly the internal market. In the pilot test, the model struggles with estimating large infrequent outliers, which are overwhelmingly geminated from a single market outside of Europe(China). When these outliers are removed the model generates estimates that are broadly

consistent with those generated by the firm. The trend in the models excluding China are boadly consistent with the trends observed from the firm's approach (general decline from 2006-2008, rise during 2009-2011), but year-to year the RAND model deviates from the firm's trend. Without information from the firm regarding the statistical uncertainty in their estimate, it is not possible to know if these year-to-year deviations are statistically meaningful but give pause regarding the ability of the model to reliably predict short term fluctuations in counterfeiting. A more thorough and careful assessment of the model, which would include data from additional firms, other products, and a longer time period, is required before it can be determined if the methodology reliably predicts long or short-term fluctuations. Also, it would be preferable to compare the results of the RAND model to observations that are exclusively based on an industry gold standard, such as mystery shopping. In the current application, the firm used a hybrid approach involving mystery shopping and modelling for generating estimates of counterfeiting that makes it difficult to ascertain the extent to which deviations in the firm data reflect real differences or some modeling variability.

Unfortunately, the preliminary assessment of the empirical model was substantially hampered because we were only able to complete a pilot test with one industry partner. The difficulty to recruit industry partners for data collection is in itself a shortcoming of the current approach, which will be discussed in further detail. Therefore, the evidence is incomplete and more piloting is needed to draw conclusions on the actual levels or trends in IPR infringements. Nonetheless there are a number of benefits associated with using this approach, should it be proven effective with additional data.

We therefore conclude based on the consistent evidence in the long run trends and statistical overlap of our level estimates and the firm's estimates in models excluding statistical outliers, that the RAND model has promise and should be more thoroughly tested and refined. The inability of the model to perform as well with outliers geminating largely from a single county is widely viewed as a major supplier of counterfeits is something that should be taken seriously, but should not condemn this approach until further testing of the model is undertaken for other firms and products (and compared to other firm estimates of counterfeit). It may be that the level of counterfeiting is so different for this single country an entirely different approach needs to be taken for it than from those countries that are generally smaller producers of counterfeits.

What are the benefits of this methodology in comparison to alternatives?

The RAND method has a number of advantages over and above approaches that have been applied in the past:

- **Cost-effectiveness.** The proposed methodology can be implemented at relatively low cost vis-à-vis other industry gold standard methods such as mystery shopping. It provides an economically feasible tool for government and regulatory agencies that need to monitor trends in counterfeiting or evaluate the effectiveness of alternative policies and interventions. For firms, this approach provides an alternative cost-effective means for filling in data gaps in other markets where the gold standard is not applied and does so in a way that is not sensitive to selection issues that can bias estimates using extrapolation methods from gold standard samples.
- **Flexibility.** The approach is relatively flexible and can be modified to meet unique aspects of specific products, firms or industries while still generating aggregated output that can be generalised across products, firms and industries to generate regional market or global

estimates of the level of counterfeiting. The flexibility comes about because of the two stage estimation process. In the first stage, a firm interested in understanding its own deviations from forecasts can customize the information in their first stage to be as firm- or product-specific as they like. In the second stage, the method is adaptable to the specificities of IPR infringements in market environments and for products.

- **Comparability.** The method enables a systematic comparison of counterfeiting effects across firms operating within the same market or in markets for similar products. This is because the same model is applied across firms, and hence any general market error that might exist in estimating counterfeiting more generally will not influence the relative effects of counterfeiting of one firm vis-à-vis other firms.

- **Replicability.** One of the main benefits of the method is its ability to be replicated for multiple products, in multiple countries and in consecutive years. Whilst the methodology, and its components – such as the second stage indicators – may be subject to change over time, it would be fairly straightforward to update the estimates retrospectively which would maintain comparability of the results over time. If the method will be improved or adjusted in the following years, the marginal extra costs of running the model retrospectively for preceding years are relatively low.

What are the challenges and limitations of this methodology; can they be tackled, and if so, how?

While preliminary evidence suggests that the RAND method does a good job at tracking the general trends reported by mystery shopping when China is excluded, a number of challenges remain. These need to be addressed or taken into account before the RAND method can be applied more broadly.

Challenges with using forecast data. The applicability of the method depends on the availability and quality of firm forecast data. There are various reasons why collecting firm forecast data may be difficult. There may be divergent business models (e.g. box office, DVD sales or broadcasting) in which sales are measured in different units (tickets, DVDs or broadcasting rights). Another concern relates to the extent to which firms incorporate counterfeiting into their forecast and whether or how this can be removed for use in a model. And related to this, some firms do not systematically generate forecasts but instead just use historical data to project trends going forward, which would inherently include influences of counterfeiting but in a fashion that is not discernible by the firm. Such issues represent a challenge for estimate of the level of forecasting systematically across markets, but assuming that such issues are firm specific and time persistent, they provide no threat to the ability of the methodology to predict trends or changes in trends of counterfeiting.

Challenges with obtaining forecast and actual sales data. Although forecast data seemed to be available in many instances, firms were extremely reluctant to share the data. There are a number of salient reasons for this reluctance. For example, there is concern about the potential for the disclosure of commercially sensitive data. Firms seemed to be reluctant to be the first participant in the study from a given industry. Finally, it may be difficult for firms to collate forecasts from different products, as the forecasts may be conducted in a decentralised manner, at national or regional market level. Concerns such as these arose with the pilot firm which we worked with as well, but were easily resolved through direct communication and education on the need for specific information. One challenge that was raised by nonparticipants is the extent to which

firms may try to manipulate their forecasting error data before submitting them to be included in our model so as to influence estimates of the size of the market. While it is true that such strategic behaviour is possible, analytic diagnostics are available that could lead to its detection if the model is implemented for all targeted products within a sector. More importantly, such biases would not influence the reliability of the model in projecting trends in counterfeiting on the long run, provided that firms were persistent in their over-reporting over time.

Industry specific concerns. Any approach attempting to generate estimates of IPR infringements in a systematic way across multiple firms and industries is going to have to necessarily aggregate measurement issues to a level that will be far less precise and meaningful than if the assessment were being done for a single firm or industry. Some industries have specific characteristics that require serious consideration. Addressing these set of challenges directly is complicated and is likely impossible without actually working with the data, but the flexibility of the model suggests they may not be insurmountable. Estimates from the second stage model may, for example, be best obtained on an industry-by-industry basis, enabling for differential inclusion of specific second stage variables. Such an approach is feasible with this model as the aggregation of "units" counterfeited by market is done after estimation of the second stage model.

The applicability to unauthorised access to protected content (UUPC). On theoretical grounds, we do not reject the possibility that our methodology might offer sensible insights on the extent of UUPC. However, from an empirical perspective, UUPC industries and particularly those involving on-line content, have a number of specificities that may complicate the applicability of the model. While we have received some input on how to tailor our model, we have not been able to test it with actual data. Therefore, it is relevant to highlight the concerns and limitations, but it is too early to dismiss the RAND methodology for UUPC altogether.

What are the next steps that need to be taken in order to assist the European Commission in its ambition to measure the development of IPR infringements in the internal market on an annual basis?

The methodology described above is a first step towards developing a system to monitor trends of IPR infringements in the internal market. Prior to implementation, the feasibility and reasonableness of this approach will need to be tested and demonstrated across multiple firms and industries. For this to happen, a number of steps must be taken next.

Build trust and buy-in from the industry. A critical next step necessary to make any further progress on developing the methodology is to build trust and create buy-in from key industry leaders. We are actively engaging academic leaders to provide their perspectives on the approach. Furthermore, we encourage stakeholders to engage in discussions about the applicability of the approach to their markets. The fact that there are weaknesses in the approach is, by itself, not a limiting reason to stop further exploration of the method. Creating buy-in may require publication of non-technical explanations or presentations to the policy community and stakeholders.

Continued development of the methodology. The utility of the method for firms and policymakers can only be understood through its empirical testing using real world data from multiple firms. This should be a priority. It will be important to confirm the proof of concept by extending the pilot work in the near future with a selection of firms representing a broad range of products, including those related to online UUPC. Assuming that a core set of variables is found to be consistently useful for predicting unexplained forecasting error, then efforts can be

broadened to assess the reliability of the approach in more competitive markets. Statistical models can relatively easily account for unique factors that are time persistent by product line or firm using a statistical tool called fixed effects. Some of these factors may be easily addressed through statistical modelling rather than complicated data gathering tasks.

Possibility to tailor model to sector specificities. A key strength of the RAND methodology proposed is its flexibility to handle contemporaneously unique industry-, firm- and market-level factors. By extending a pilot to multiple product groups and industries, it will also be possible to consider the extent to which unique industry characteristics might impede the implementation of this approach. Much of the discussion has focused on the identification of common aggregate measures of IPR infringements at national level. But the RAND model could also be applied on a sector-by-sector basis, which would enable a more explicit consideration of sector-specific attributes.

Facilitating data delivery. The process involved in identifying the data required for this pilot, collating them in the correct format from the firm, and properly structuring it for estimation in the model has been relatively time consuming and cumbersome both for the researchers and firm representatives involved. There are several steps that can be undertaken to facilitate and accelerate this process: 1) A research team member needs to spend time with the firm to explain the approach, understand their forecasts and sales trends and how data describing those trends are captured by the firm; 2) Robust provisions, including signed data use agreements, are required for data protection; 3) A standardised template for data submission should be prepared to facilitate the delivery of data in a systematic way across all firms.

Abbreviations

ACTA	Anti-Counterfeiting Trade Agreement
BSA	Business Software Alliance
C&P	Counterfeiting and Piracy
CD	Compact Disc
CEBR	Centre for Economic and Business Research
DVD	Digital Versatile Disc
FDI	Foreign Direct Investment
GAO	US Government Accountability Office
GNI	Gross National Income
IFPI	International Federation of the Phonographic Industry
INTA	International Trademark Association
IPR	Intellectual Property Right
NBER	National Bureau of Economic Research
OECD	Organisation for Economic Cooperation and Development
OHIM	Office for Harmonization in the Internal Market
P2P	Peer-to-Peer
PC	Personal Computer
R&D	Research and Development
SME	Small and medium– sized enterprise
TRIPS	International Agreement on Trade-Related Aspects of Intellectual Property Rights
UUPC	Unauthorised Use of Protected Content
WIPO	World Intellectual Property Organization

Acknowledgements

Introduction

'Counterfeiting' can be understood as the production of fake or forged goods, while the label 'piracy' is often used for unauthorised use of protected content (UUPC). Both counterfeiting and piracy are commonly-known types of violation of intellectual property rights (IPRs). Examples of counterfeited products include medicines, luxury goods and spare parts for vehicles and aircraft, while popular forms of piracy in this context include illegal downloading and copying of films, music and software.

1.1 Why this study is relevant

The trends of globalisation and the integration of markets in recent decades ahead of corresponding integration of IP law may have facilitated a rapid spread of IPR violations. Furthermore, the increasingly widespread access to computers, Internet and other technological developments help those involved in these illicit businesses to duplicate designs, labels, logos, packaging and documentation with speed, accuracy and relative anonymity (Treverton *et al.,* 2009). There has been much debate about the extent of counterfeiting and UUPC, as well as about their impacts.

Some argue that 'victims' of counterfeiting and piracy may face considerable economic, health and safety impacts. Many of these will impact the IPR holders. For example, company profits or brand names may be compromised by the existence of fake and pirated products. The willingness of firms to invest in R&D and to innovate might be affected. Governments could lose tax revenue. And finally, consumers may knowingly or unknowingly spend a significant amount of money on a fake product that does not perform as expected. And, as a consequence, their health or safety may be at stake.

Others have pointed to paradoxical positive effects: the availability of counterfeit luxury goods to poorer markets primes those markets for later development by the victim companies; counterfeit workshops and businesses act as nurseries for manufacture and commerce in developing economies and pirate music and film act as a marketing tool for producers. However, no such paradoxical effects are posited for forms of counterfeiting and piracy that create a significant threat of harm (aircraft and vehicle parts, drugs etc.) For that reason, classifying IPR violations according to the nature and significance of the threat posed has been advocated.

Without objective and reliable estimates of the extent of IPR violations it is difficult to debate these claims. Many attempts have been made to develop and use various methodologies to estimate the number of illegal products traded, the value of these illegal products, the value of real products these pirated or counterfeit goods replaced (lost legitimate revenue) and the social and

economic impacts of counterfeiting and piracy. Most of these methodologies, however, have been developed to the estimate the size of *specific* markets for these illegal products, for example illegal software or tobacco. A recent publication of the US Government Accountability Office (2010) justifies this approach: 'Because of the significant differences in types of counterfeited and pirated goods and industries involved, no single method can be used to develop estimates.'

Not only have existing approaches been inadequate for measuring counterfeiting and UUPC of a range of products simultaneously; these approaches also suffer from a number of important methodological weaknesses (described in greater detail in subsequent chapters), even when they focus on specific goods or industries. This means that the resulting estimates must be heavily caveated and qualified, putting into question the extent to which they are useful to governments and firms trying to understand and tackle the phenomenon.

This document presents the final deliverable of a study to develop a methodology for assessing the scale and impact of counterfeiting and piracy in the European Union. The report presents findings from the three main tasks of the study:

1) an extensive review of literature and, specifically, of the methodologies that have been used to estimate the magnitude of counterfeit and piracy;
2) preliminary steps in the development of the estimation methodology;
3) pilot-testing this methodology with empirical data.

1.2 Objectives of this study

Given the intensity of the debate around IPRs, echoed in the controversy around the ratification of the Anti-Counterfeiting Trade Agreement (ACTA) for instance, and the potentially significant economic and other interests that play a role, an objective and evidence-based approach towards measuring the scale and impact of the phenomenon has become more important than ever. This study was set up to assist the Internal Market and Services Directorate General (DG MARKT) of the European Commission and the European Observatory on Infringements of Intellectual Property Rights in the development of a methodology that would quantify the scope, scale and impact of IPR infringements on the European economy. This study is the first stage in an attempt to continuously assess the problem and to develop evidence-based policies in the area of IPRs.

In order to address this objective, we have formulated a number of research questions that this report aims to answer:

- What can we learn from previous efforts about the drivers and impacts of IPR infringements?

- What are the strengths and weaknesses of existing methodologies that have been applied to measure the scope, size and impact of IPR infringements?

- What does this mean for the development of a methodology to be applied by the European Commission to estimate the scale of IPR infringements in the internal market?

- What would be the characteristics and data requirements of such a methodology?

- What does application of this methodology teach us about the scale and impacts of IPR infringements in the internal market?

- What are the benefits of this methodology in comparison to alternatives?

- What are the challenges and limitations of this methodology? Can they be tackled and, if so, how?
- What are the next steps that need to be taken in order to assist the European Commission in its ambition to measure the development of IPR infringements in the internal market on an annual basis?

As the remainder of this report will clarify, most of these questions have been addressed. However, several challenges remain. Based on the lessons from previous studies, we decided to focus our attention on the impact of IPR infringements on the sales of legitimate goods and products. We have developed a methodology that can be used to monitor the trends in this area.

While this is only one of the potential impacts of infringements, the study acknowledges that measurement of other impacts (e.g. on innovation, growth, competitiveness, creativity, culture, public health and safety, employment, environment, tax revenues and crime) cannot advance unless the measurement of infringement itself has reached a scientifically satisfactory stage.

In this report we offer the 'building blocks' for a methodology that is consistent, robust, feasible and reliable in measuring the size of counterfeiting and UUPC. In a next stage it would then be sensible to work on improving the methodologies that are currently applied to the study of the broader impacts of infringement.

1.3   Structure of this report

In order to address the objectives outlined above, we have structured the report as follows.

First of all, we provide a brief overview of the definitions relevant to this report. Chapter 2 defines IPRs and discusses their different types. More importantly, we explain what we mean by infringements of IPRs and distinguish between counterfeiting and unauthorised use of protected content.

Before attempting to measure the scale of IPR infringements and their impacts, it is important to understand the drivers of IPR violations from both the demand and the supply side. These factors help to explain the characteristics of products that are subject to IPR counterfeiting or piracy. Chapter 3 reviews the available literature on these drivers. This chapter also presents an overview of the variety of estimates available in the literature of the magnitude of IPR infringements as well as the breadth and scale of the impacts.

Our proposal for the development of an approach to measure IPR violations in the internal market is based on an extensive review of the data and methodologies that we have been able to identify. Chapter 4 presents a synthesis of these findings. More information on individual sources in the literature can be found in Appendix A.

Based on the findings from this extensive review, we present a theoretical framework in Chapter 5 and propose a methodology for measuring trends in the lost revenues in markets of counterfeited products due to IPR violations. This methodology has been piloted with empirical data from a single firm in the technology industry demonstrating the potential value of this approach for the European Commission. Whilst we and many of those who have reviewed this work believe that this methodology presents a most promising approach to the problem, there are still various limitations and caveats involved with the method. These are summarised and, where possible, addressed in Chapter 6.

Chapter 7 discusses the applicability of the methodology to UUPC, which (due to the characteristics of these products and markets) may be challenging. We suggest that an approach to measure the extent of UUPC in the internal market, potentially based on our proposed methodology, needs to be investigated in more detail.

Finally, Chapter 8 offers a number of suggestions for next steps to take this approach to the next level. Before we recommend implementing a measurement system based on the proposed methodology, it will need to be pilot-tested in on more products from different sectors. Chapter 8 lays out these next steps in more detail.

Defining IPRs and their infringement

In this chapter we provide a brief overview of the definitions relevant to this report. It is important to understand the scope of the study, and therefore we define IPRs and discuss some their common types. More importantly, we explain what we mean by infringements of IPRs and distinguish between counterfeiting and unauthorised use of protected content (UUPC).

2.1 What is an intellectual property right?

Although conceptually and legally different concepts, both counterfeiting and UUPC constitute illicit activities linked to IPR infringement. IPRs enable creators, businesses and investors to protect their tangible and intangible products by preventing unauthorised exploitation of their goods or by allowing such exploitation in return for compensation (EC DG Trade, 2011b). In that way, proportionate protection of IPR plays an important role in innovation, creativity and competitiveness, and is considered crucial for building a knowledge economy.

2.1.1 Types of IPR

IPRs are broadly divided into two main areas: copyright (in common law countries) or authors' right (in civil law countries) on the one hand, and industrial property on the other. A recent document published by the European Commission describes all the rights that are relevant to intellectual property (European Commission, 2011). Copyright and rights associated with industrial property are briefly described below.

Figure 2-1: Rights relevant to intellectual property

Adapted from European Commission (2011)

Copyright and rights related to copyright

Copyright covers a wide range of works ranging from creative works, such as books, music, films, performances and broadcasts, to technical works, such as computer programmes, games or software. Copyright applies from the moment of creation of the work and provides the author with the exclusive right to prevent third parties from using this work without authorisation. Copyright protection is time-bound and usually corresponds at least to the natural life of the creator plus 50 years after his death (OECD, 2008; TRIPS, 1994).[1] In addition to the rights of the author, copyright legislation also provides protection to auxiliaries and intermediaries who contribute to the dissemination of works. This means, for instance, that copyright also protects the rights of music producers in their CDs and digital music, and the rights of broadcasting organisations in their radio and television programmes. While protecting the rights of creators, copyright balances public and private interests by allowing reproduction of a protected work for personal and private use, or for public use in some cases (Bryce, 2009).

Property rights

The second broad area of IPRs covers industrial property rights, including trademarks, geographical indications, patents and licensing.

Trademark

A trademark is any distinctive sign that identifies goods or services produced or provided by a specific person or enterprise and distinguishes them from the goods and services of competitors. According to the TRIPS definition, 'any sign or a combination of signs' can constitute a trademark; these include 'personal names, letters, numerals, figurative elements and combinations of colours as well as any combination of such signs' (TRIPS, 1994). The owner of a valid trademark has an exclusive right to use it. In addition, others are excluded from using similar or identical marks for similar or identical products (Bryce, 2009; TRIPS, 1994).

Geographical indications

Geographical indications are forms of identification of names and symbols which identify products as originating in a particular region or locality. As geographical identifications identify a geographical area in which one or several producers are located, there is no owner of geographical

---

[1] was extended to 70 years within the European Union through Directive 2011/77/EU, amending Directive 93/98/EEC.

identification and all enterprises in a particular region have the right to use the indication for products originating in that region. Protection of geographical indications consists of prevention of unauthorised persons from using them either because products do not originate from the geographical place indicated or because products do not comply with the prescribed quality standards (Bryce, 2009; EC DG Trade, 2011a).

Patents

According to the International Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) definition, patents 'shall be available for any inventions, whether products or processes, in all fields of technology, provided that they are new, involve an inventive step and are capable of industrial application' (TRIPS, 1994). A patent is a document issued by a national or regional patent office which describes an invention and creates a temporary legal situation in which the patented invention can be made, used, offered for sale only by the patent holder or with the patent holder's authorisation. As a result, the exclusive right of a patent owner has two main applications: protection against infringement and the possibility of assigning or licensing the right. A patent right is valid for at least 20 years from the date of filing an application for a patent. This period of exclusivity is granted, however, on condition that a patent holder must disclose information about the patented product to the public so as to stimulate further research and innovation. Patent rights are geographically bound and patent application must be made in every jurisdiction in which a prospective patent holder wants to protect the product, although some of the burden of multiple applications is alleviated by a more centralised application procedure through the Patent Co-operation Treaty (Bryce, 2009; OECD, 2008; TRIPS, 1994).

2.1.2 The main objectives of IPRs

IPRs serve both economic and social functions. In its economic function, an IPR provides the rightful holder a competitive advantage by preventing unauthorised exploitation by third parties (EC DG Trade, 2011c). A company that has protected its products or processes by IPR can derive revenues from direct and indirect exploitation of these rights. Indirect exploitation by third parties under licensing contracts sometimes can exceed the profits from direct exploitation, in particular for small and medium enterprises (SMEs), universities and other public research centres that usually do not have any direct exploitation activities.

The wider societal benefits of IPRs include access to up-to-date technical information, facilitation of technology transfer and protection of safety standards. As most patent information is usually published 18 months after the first filing, patent information is a way of rapidly disseminating up-to-date technical information. In addition, because patent descriptions tend to provide accurate information about technical specification of a product or process, they facilitate technology transfer and similar agreements. This trade facilitation function means that 'patents have sometimes been considered as the currency of the knowledge-based economy' (EC DG Trade, 2011c). From the customer's perspective, IPRs provide guarantees regarding the intrinsic quality and safety of products, and compliance with applicable safety standards (EC DG Trade, 2011c).

Despite benefits deriving from IPRs, the protection of inventions is challenging. First, the single market for patents remains incomplete and protection of intellectual property and validation of patents by country bring significant costs for companies. Second, the lack of a unified and specialised patent litigation system means that companies that want to enforce or challenge patents have to take a costly and legally risky route.

2.2   Infringements of intellectual property rights

The production, distribution, storage and sales of products (such as trademarked or patented products) by non-holders of an IPR is an infringement of IPR legislation. These are commonly referred to as counterfeit and pirated (C&P) products, and they may include music, film, software, medicines, fertilisers, aircraft and car parts, luxury goods (such as bags and watches) and a wide range of other goods. As explained above, IPR violations are thought to impact industry and government interests. As a consequence, industry initiatives to combat illicit activities are already taking place, and the response from international and national organisations is intensifying.

Industry efforts, pursued at the firm- and sector-level, as well as cross-sector initiatives, focus on four main areas. First, industries conduct research and collect information about counterfeiting and UUPC practices in their sectors; these data are used to develop public awareness about illicit products and develop counter-measures. Second, legitimate goods producers undertake various steps to make their products more difficult to copy and counterfeit, for example through improvements in authentication and track-and-trace technologies. Third, industry representatives are involved in supporting government efforts to combat counterfeit and UUPC; this takes the form of training and awareness-raising programmes delivered to police, prosecutors, customs officials and enforcement personnel in the producers' own country as well as in third-party countries. Fourth, industry takes legal action and pursues violators of IPR through courts (OECD, 2008).

However, in line with some of the paradoxical effects and characteristics of counterfeiting and piracy some industries have also taken steps to reach a more rounded understanding of the way these phenomena impact their business. This has led to a revolution in the way the music business operates; a degree of positive interaction between counterfeiters and brand owners in the fashion industry and the development of valuable branded counter counterfeit systems marketed by drug companies.

Recent years have also seen several initiatives to enhance international co-operation to reduce trade in counterfeited and pirated products, through improvements in the effectiveness of intellectual property policies and programmes and closer international collaboration of stakeholders. This includes initiatives led by the World Trade Organization, the World Intellectual Property Organization, the World Customs Organization, the World Health Organization and others.

2.3   Defining IPR infringements

While the scope of this study focuses on IPR infringements in general, the two concepts that get most airplay are 'counterfeiting' and 'piracy'. In this report, 'piracy' is referred to as 'unauthorised use of protected content' (UUPC) except when we are citing literature and other sources that specifically use the term 'piracy' – more on this below. TRIPS (1994) defines counterfeiting as follows:

> 'counterfeit trademark goods' shall mean any goods, including packaging, bearing without authorization a trademark which is identical to the trademark validly registered in respect of such goods, or which cannot be distinguished in its essential aspects from such a trademark, and which

thereby infringes the rights of the owner of the trademark in question under the law of the country of importation.

This definition essentially sets the standard for how counterfeiting is understood in research and analysis. However, when we refer to counterfeit products we do not limit ourselves to trademark violations. Hence we generally adhere to the broad definition introduced by the OECD (1998):

> … the term 'counterfeiting' is used in its broadest sense and encompasses any manufacturing which so closely imitates the appearance of the product of another to mislead a consumer that it is the product of another. Hence, it may include trademark infringing goods, as well as copyright infringements. The concept also includes copying of packaging, labelling and any other significant features of the product.

While in this definition copyright infringements are considered counterfeiting as well, we argue that these violations clearly distinguish themselves from counterfeiting. Piracy is a popular term for such infringements.

The term 'piracy' captures the intrinsic quality of this process: the goods misappropriated in a traditional act of piracy at sea were genuine, but the misappropriation was concealed and thereafter the route to market was conventional, except that a degree of complicity was usually necessary from middlemen. The implication is that protection from unauthorised access needs to operate, if possible, before the misappropriated intellectual property is concealed or absorbed into a regular and licit marketing process, where and when it becomes indistinguishable from the legal equivalent.

Although the term is commonly used in the literature as well as in popular media, it is not uncontroversial and can be subject to multiple interpretations. Therefore we use the term 'unauthorised use of protected content' (UUPC). Products characterised by such unauthorised use are mostly (but not always)[2] in violation of copyright. This aspect of authorisation is also reflected in the definition provided by TRIPS (1994):

> 'pirated copyright goods' shall mean any goods which are copies made without the consent of the right holder or person duly authorized by the right holder in the country of production and which are made directly or indirectly from an article where the making of that copy would have constituted an infringement of a copyright or a related right under the law of the country of importation.

In its report on digital piracy, the OECD (2008) uses the definition of piracy suggested by TRIPS, but focuses on copyright infringements that cover 'only Internet and direct computer to computer transfers, LAN file sharing, mobile phone piracy and so on". We believe that the term 'digital' is confusing in this context, as tangible goods such as CDs, DVDs, flash drives, etc. are also digital media. Instead of 'digital', therefore, we will use the term 'online' when referring to UUPC of non-tangible goods available through the Internet, file-sharing and so on.

---

[2] For example unauthorised streaming of broadcasting (e.g. sports or music events) is not necessarily in violation of copyright. It is UUPC, however, as broadcasting rights and contracts limit who can show footage of the event.

CHAPTER 3   **The drivers, scale and impacts of IPR infringements**

Research about why people produce and consume counterfeit and pirated products has been a fixture of the counterfeiting and UUPC literature for decades. Much of the most recent research on this issue has focused on what is labelled as illegal downloading of music, film and software, although studies have been conducted that explore the drivers of other types of unauthorised use of content and counterfeiting as well.

In this chapter we review some of the key insights from the literature on the drivers of IPR infringements and their potential impacts.

3.1   **The drivers of IPR infringements**

In this section we summarise some of the key insights from a review of the literature.[3] We divide this broadly into determinants of the supply of and of the demand for counterfeited and pirated products. Given the fast pace of change in the technology that enables UUPC and counterfeiting, we focus on research carried out in the last decade, although older studies are also cited when they offer important or unique perspectives and insights.

3.1.1   **Supply-side determinants**

While estimations of the scale of counterfeiting and UUPC are plagued by limitations, all indications are that the incidence of counterfeiting and UUPC is increasing rapidly. On the supply side, many reasons have been posited for this. Wall and Large (2010, p.1097) have argued:

> The wide availability of digital and networked technologies now means that information about products and their production is more easily accessible and that advances in industrial capability now make it much easier to commission, manufacture and sell counterfeit goods globally. The globalization of world trade has broadened consumer desires by making available a wider range of goods and this ever-increasing demand for luxury goods makes them attractive to counterfeiters. Furthermore, the creation of free markets also assists in the sales process, because once goods have entered the European free market, they can freely circulate across the borders of its member states. Also attractive to counterfeiters is the fact that in many countries, *de jure* or *de facto* penalties for counterfeiting are relatively low and in some cases non-existent … Even countries that possess severe legal penalties for counterfeiting, such as Italy, which has criminalized the purchase of counterfeit goods, have minimal enforcement. Finally, and of considerable concern in

---

[3] For a more extensive discussion of the drivers of counterfeiting and piracy, see OECD (2008).

governmental circles, is that counterfeiting is becoming increasingly attractive to organized crime because of the high profits and lower levels of risk than more traditional criminal activity.

The drivers presented by Wall and Large (2010) refer primarily to macro-level incentives and developments that create an environment conducive or favourable to counterfeiting and UUPC in general. Other studies have come to broadly similar conclusions. A recent study by the Swiss Federal Institute of Technology and the University of St Gallen (Staake, 2009) argued that growing production capabilities, weak enforcement of IPRs, strong demand within emerging markets, the growing importance of global brands (especially for luxury products) and the global integration of trade are all key contributors to the growth and widening scope of the counterfeit trade.

A few studies and observers have examined the supply-side determinants of particular counterfeit or pirated goods. For instance, according to Bale (2005) it is possible that as counterfeit medicines can be made relatively cheaply and are likely to be at least as profitable as illicit drugs, pharmaceutical products became a particular target for counterfeiters. In addition, the fact that end-users have little knowledge of the product (it is a 'credence good') means that they are particularly vulnerable to this crime.

The Business Software Alliance has also shed light on the supply-side determinants of software piracy. It identifies three drivers as having a positive correlation with UUPC (BSA-IDC, 2010):

- Rapid growth of the consumer PC market;

- Activity in the base of older computers where unlicensed software is more prevalent; and

- Increasing sophistication of online criminals leveraging the Internet and other new means of distribution.

It is important to note that in the specific case of unauthorised use of protected online content, the distinction between supply and demand of pirated products is much more blurred than in the case of counterfeiting. This is because, particularly in peer-to-peer (P2P) networks,[4] many consumers of these products are at the same time unauthorized publishers of software, music, films and other products.

3.1.2 Demand-side determinants

A key characteristic of the demand for counterfeits is deception. As Grossman and Shapiro described it in their seminal work (1988), consumers can either be deceived by counterfeit producers and think that they are purchasing the original product, or willingly purchase a fake in order to lead others to believe they own the original. In the former situation, counterfeit goods compete on the primary market, i.e. they infiltrate the market of the genuine good. In the latter situation, when the consumer actively seeks the infringing good, a secondary market is established where an explicit demand for lower quality and price is met by unauthorised producers. While on the primary market the buyer has imperfect information as to the real quality of the good, which the producer exploits to erode consumer welfare and earn extra profits, on the secondary market the buyer exploits the imperfect information of his peers (e.g. in the case of clothing) or of the

---

[4] A peer-to-peer network is computer network of computers typically connected over the Internet, in which each computer can act as a client or server for the others. P2P networks can be used for sharing content such as audio, video, data or anything in digital format.

authorities (e.g. in the case of unauthorised copies of music or movies) in order to derive utility from the consumption of lower-quality and lower-price goods.

It is therefore important to consider how much utility the consumer derives from the infringing good. On the primary market, the question is less relevant as the consumer derives the same utility he would have derived from the original good as long as he does not realise that he has been deceived (OECD, 2008, p.45). On the secondary market, a relationship of substitutability is established between the original good and the counterfeit or UUPC. This relationship is referred to as the substitution rate (ibid.), which defines the rate at which the consumer is willing to substitute the original good for an infringing one.[5] On the primary market the substitution rate is 1 (i.e. the counterfeit product is a perfect substitute if the consumer is deceived), while on the secondary market it is lower because the consumer recognises the counterfeit and willingly purchases it at a lower price (expecting lower utility from it). This implies that on the secondary market the purchase of one counterfeit good may not correspond to the foregone sale of one original good. Whether the purchase of a counterfeit or pirated good in the secondary market has a negative or positive effect on brand equity depends on the impact of this demand on the demand for the genuine products. Therefore the substitution rate provides a convenient way to express the displacement of sales of the original product caused by counterfeiting or UUPC.

A growing body of research has examined the determinants of the demand for products in violation of IPRs. Broadly, studies have examined the extent to which macro-level factors (social, economic, legal, political and cultural) and micro-level factors (attitudes, the price of a product, etc.) affect or are associated with the rate of UUPC and counterfeiting.

A significant fraction of the literature reviewed focuses specifically on the demand-side determinants of online UUPC. At the macro-level, studies tend to find that the better the socio-economic indicators, such as educational investment and GDP, the lower the rate of UUPC. For instance, a recent study examined the correlation between the 'business environment' in a country and the country's degree of UUPC, specifically software piracy (Yang, 2007). A sample of 76 countries was divided into 'high- and low-piracy countries' (i.e. with piracy rates above or below 60%). Statistical analysis of the link between piracy and various indicators of the business environment shows that many of these indicators affect the incidence of piracy. For instance, Gross National Income (GNI) per capita appears to be inversely correlated with piracy, but only in high-piracy countries. That is, in most developed countries it appears that people who buy pirated products 'are driven by factors other than their income' (*ibid*, p.138). This finding echoes those of other studies, which suggest that GNI or GDP per capita are negatively associated with piracy, and that they account for much of the variation in piracy rates between countries (e.g. Burke, 1996; Depken and Simmons, 2004; Ronkainen and Guerrero-Cusmano, 2001). Similarly, the study found that education expenditure has a negative impact on piracy in high-piracy countries, but no impact in low-piracy ones. In low-piracy countries, individualism, defined as the degree of emphasis on individual rights and freedoms, also appears to be negatively associated with low-piracy rates, a finding also echoed in other studies (for example Al-Jabri and Abdul-Gader, 1997; Yang, 2005; Shore *et al.,* 2001). This may be because in individualist

---

[5] While this informal discussion provides an intuitive way to think about consumer behaviour, it is important to highlight that economics defines the substitution rate in marginal terms, i.e. with respect to infinitesimally small changes. The marginal rate of substitution depends on the relative quantities of the two goods and on individual preferences.

countries software and other types of piracy are strongly perceived as a violation of the IPRs of an individual (i.e. the rights holder).

Whether broadband penetration and Internet access are associated with UUPC is a debated element of the literature. The International Federation of the Phonographic Industry (IFPI, 2009) implicitly suggests that broadband penetration is not a determinant of file-sharing or that there is an inverse relationship, since as broadband penetration increased, file-sharing decreased. On the other hand, while the methodology used by Business Software Association is not transparent enough to be certain, their estimates suggests that increasing broadband increases file-sharing.

At the micro-level, most research on consumption drivers has been conducted through surveys and has focused on attitudes and intention to buy or experience of buying or using counterfeit and pirated products. One study of the UK market for luxury counterfeit goods suggests that consumers are increasingly willing to buy these knowingly as a result of improvements in the actual and perceived quality of the counterfeit products, and of increasing social acceptability of the consumption of fakes (Ledbury Research, 2007).

Other micro-level drivers have also been identified. In relation to music piracy in particular, one study of young people in Taiwan suggests that the fact that pirated music tends to be of good quality is a strong driver of subsequent UUPC (Chiou *et al.*, 2005). The growing social acceptability of UUPC and perceived low risk of penalties also affect piracy intentions (ibid.). While social acceptability of UUPC and counterfeiting was also found to affect attitudes towards those activities in Europe (Van Barneveld, 2010), other factors seem to be less homogeneous cross-culturally. For example, a study of American and Indian students found that ethical considerations may play different roles in determining attitudes to piracy across different cultures (Gopal and Sanders, 1998). An experimental study in Hong Kong and Las Vegas shows that the Hong Kong subjects were more likely to purchase counterfeit goods than the Las Vegas subjects, although both respond to changes in the price of the authentic good and the expected penalty for purchasing counterfeit goods (Harvey and Walls, 2003).

A survey in the European Union conducted by Eurobarometer asked the question: 'When is it OK to buy counterfeit products?' (Van Barneveld, 2010). It is not clear from the document reviewed whether the answer 'never' was available to survey respondents. However, 33% of respondents agreed or agreed very much that it is acceptable to buy counterfeits when the price for the original is too high, 27% agreed or agreed very much that it is OK when the original is not available, 25% agreed or agreed very much that it is OK when quality does not matter and 21% agreed or agreed very much that it is OK when it concerns a luxury.

3.2 The scale of IPR infringements

The measurement of the scale of counterfeiting and UUPC has been the subject of much discussion in the literature. Several attempts have been made to develop and use various methodologies to estimate the number and value of infringing products. Most of the methodologies used to measure these have been developed for specific products; motivating this tendency, a recent publication of the U.S. Government Accountability Office (GAO, 2010) argues   that "because of the significant differences in types of counterfeited and pirated goods and industries involved, no single method can be used to develop estimates." Following Olsen

(2005), the methods used can be classified as enforcement-based (exploiting information on seizures or legal actions), surveys of consumers, producers and distributors, sampling (mystery shopping) exercises, and economic and econometric models. Published examples of the application of these methods are reviewed in detail in Appendix A. The remainder of this section presents an overview of the existing estimates of the scale of counterfeiting and UUPC. The aim is to describe the range of point estimates and the variation in the metrics used, as summarized by Table 3-1, rather than comprehensively review them. While no single estimate can be regarded as definitive, the indication that emerges from this array of studies is that the phenomena of counterfeiting and UUPC are sizeable. Table 3-1 provides a non-exhaustive list of studies that have attempted to measure the magnitude of IPR infringements for a variety of products.

3.2.1 Overview of existing research on the scale of counterfeiting

While the importance of counterfeiting has been recognized by the private and public sectors for quite some time, the intrinsic difficulty in building a practical and empirical understanding of the phenomenon is reflected in the relative scarcity of original estimates of its magnitude in the literature. This sub-section focuses on estimates of the scale of counterfeiting (or counterfeiting and UUPC when no separate figures are provided), while sub-section 3.3.2 discusses UUPC specifically.

A number of studies exist in the non-academic realm that try to quantify the magnitude of counterfeiting across a broad variety of products. Some of these studies rely on relatively simple arithmetic: the general structure of the estimation process is built off of information on the number of infringements, the substitution rate (i.e. the number of legitimate goods that would have been bought in absence of the counterfeit), and the price at which such units would have been sold. The product of these three factors represents the value of sales lost by the IPR holder and is typically used as a measure of the value of counterfeiting. For example, following this type of structure the Centre for Economics and Business Research (2000) estimated annual lost revenues in the European Union at Euro 7,581 million in clothing and footwear, Euro 3,731 million in toys and sports equipment, Euro 3,017 million in perfumes and cosmetics and Euro 1,554 million in pharmaceuticals; a few years later, Allen Consulting (2003) valued lost sales at $ 131.7 million in the toy industry in Australia. As discussed in the Government Accountability Office report (2010), the large variation in the values that can be attached to the same quantity of counterfeits testifies to the difficulty of quantifying lost revenues: as an example, the U.S. Customs and Border Protection "seized a shipment of counterfeit sunglasses from China and reported an estimated total domestic value at $12,146 and a manufacturer's suggested retail price at $7.9 million" (ibid.). In fact, an alternative way of reporting the estimated scale of counterfeiting is in percentage terms of the total market size. For example, KPMG (2008) estimates counterfeiting in the United Arab Emirates and reports levels of 10-15% for food and beverage, 12.5% for automobile spare parts, 8-10% for cosmetics, less than 5% for cigarettes, 3-5% for household products and less than 0.1% for pharmaceuticals.

Another way in which estimates of the prevalence of counterfeiting have been produced is through surveys of consumers and producers. On the consumer side, the Gallup survey is the largest effort to date in terms of geographic coverage and estimated that globally 25% of consumers purchase counterfeits, although with large cross-country variation (e.g. 14.8% in Estonia versus 38.4 in Russia) (The Gallup Organization, 2007). On the producer side, an example is represented by the study conducted by the Swiss Federal Institute of Intellectual Property (SFIIP, 2004), which found that 54% of companies reported to be affected by IPR

infringements. An earlier study by the International Trademark Association (INTA, 1998) combined producer surveys with econometric analyses and estimated a 22% loss in sales by participating companies in 1995 due to trademark infringement and counterfeiting.

The first study attempting to get at the global market for counterfeiting and piracy across sectors, including only internationally traded goods, was an estimate by OECD (2008) of $200 billion in 2005, updated to $250 billion for 2007 (OECD, 2009). Frontier Economics (2011) built off of this analysis by using the OECD method for internationally traded goods and including domestically produced and consumed counterfeit and pirated products and digitally pirated products: the study provided a range of $455 - $650 billion for the total value of counterfeit and pirated products in 2008, with projections of $1,220 - $1,770 billion in 2015.

Finally, one alternative to the estimation strategies described above relies on so called "mystery shopping", which consists of purchases of the same specific products from a random sampling of outlets that are then sent to experts for examination to determine if the goods are authentic or counterfeit. An example of a recent study using this approach comes from the European Alliance for Access to Safe Medicines (2008), which found that 62% of medicines in their sample were substandard or counterfeit. From our discussion with experts and private sector representatives, it appears that it is not uncommon for companies to conduct this type of estimation exercises; however, the results are not publicly available.

3.2.2 Overview of existing research on the scale of unauthorised use of protected content

Alongside the heated debate on the impacts of copyright infringement on the sales of legitimate creative goods, the movie, music and software industries have made efforts in recent years to regularly produce and disseminate estimates of the scale of UUPC. Similarly to the case of counterfeiting, the prevailing metrics are a so called "piracy rate" and the value of lost sales to legitimate producers. An overview of the orders of magnitude of some of these estimates is provided in the remainder of this sub-section.

In the software industry, a 2010 Business Software Alliance study reported an estimated global software piracy rate of 43% for 2009, expressed in terms of units of pirated software installed relative to total units of software installed. This represented a significant increase from the 35% figure of 2005 (OECD 2008, p. 77). In the movie industry, a LEK Consulting study (Motion Picture Association of America, 2006) found that Motion Picture Association member companies lost $6.1 billion in revenues due to piracy in 2005. Specifically, about $2.4 billion were attributable to bootlegging, $1.4 billion to illegal copying, and $2.3 billion came from online UUPC. Finally, in the music industry, the International Federation of Phonographic Industry reported a global average piracy rate of 38% in 2006 as a share of total sales (Siwek, 2007, p. 20). At the EU-27 level, a 2010 Tera Consultants study estimated Euro 5.3 billion lost revenues in the audiovisual sectors and Euro 4.5 billion in the software industry.

As the relative size of online UUPC has grown over time, in recent years several reports attempting to estimate it have been released, introducing ad hoc metrics that relate to the very specific means of access to infringing content. For example, MarkMonitor (2011) found that the 10 media brands analyzed (in the movies/TV shows, music and software/videogames sectors) generated more than 53 billion visits a year to websites offering materials infringing their IPR. In the e-book sector, Attributor (2010) estimated the global daily demand for pirate copies at 1.5 – 3 million people. Finally, NetResult and Envisional Ltd. (2011) found an average of 197.1

websites providing illegal streaming of live football matches over a monitoring period of 12 months.

The technical aspects of these methodologies are reviewed in detail in Appendix A: suffice here to say that often the lack of clarity in fully describing the methods, assumptions and data underlying them constitutes a major barrier to an independent assessment of the statistical consistency of the results. This issue was also highlighted by earlier research efforts on the topic (e.g. OECD, 2008, p.78). In other cases, in the reports that produce the estimates reviewed above, more substantial issues remain poorly addressed from a scientific point of view: Chapter 7 mentions the important debate in the academic literature on the direction and magnitude of the effects of copyright infringements on legitimate sales. The next section turns the attention to the impacts of counterfeiting and UUPC.

3.3   The impacts of IPR infringements

There is considerable debate on the scope and magnitude of the impacts of IPR violations, as well as the mechanisms underpinning them. There are relatively few estimates of these impacts, and the robustness of these available estimates is debatable. Moreover, the literature is uneven in terms of how much attention has been paid to the different potential impacts of counterfeiting and piracy. For instance, much of the existing literature focuses on the direct impact of counterfeiting and piracy on the sales and profits of the right-holders (Olsen, 2005).

Table 3-1   Examples of types and estimates of the magnitude of IPR infringements identified in the literature

Sector	Source	Year to which the estimate refers	Countries	Metric	Estimate
Food and Beverage	KPMG (2008)	2005	United Arab Emirates	% of total market	10-15%
Automobile Spare Parts	KPMG (2008)	2005	United Arab Emirates	% of total market	12.50%
Cosmetics	KPMG (2008)	2005	United Arab Emirates	% of total market	8-10%
Cigarettes	KPMG (2008)	*	United Arab Emirates	% of total market	<5%
Household Products	KPMG (2008)	2005	United Arab Emirates	% of total market	3-5%
Pharmaceuticals	KPMG (2008)	2005	United Arab Emirates	% of total market	<0.1%
All	The Gallup Organization (2007)	2005-2006	Global (51 countries)	% consumers purchasing counterfeits	25%
All	Swiss Federal Institute of Intellectual Property (2004)	2004	Switzerland	% companies reporting being affected by IPR infringements	54%
All traded goods	OECD (2009)	2007	Global	International trade in counterfeit and pirated goods	$250 billion
All	Frontier Economics (2011)	2008	Global	Internationally traded counterfeit and pirated products	$285 billion - $360 billion
All	Frontier Economics (2011)	2008	Global	Domestically produced and consumed counterfeit and pirated products	$140 billion - $215 billion
All	Frontier Economics (2011)	2008	Global	Digitally pirated products	$30 billion - $75 billion
Pharmaceuticals	European Alliance for Access to Safe Medicines (2008)	*		% substandard or counterfeit products of total sold online	62%
Software	Business Software Alliance (2010)	2009	Global	Software piracy rate	43%
Music	International Federation of Phonographic Industry as reported in IPI (2007)	2006	Global	Piracy rate (as % of total sales)	38%

Sector	Source	Year to which the estimate refers	Countries	Metric	Estimate
Movies/TV shows, Music and Software/Videogames	Mark Monitor (2011)	*	Global	Number of visits to websites offering infringing materials	53 billion per year
E-books	Attributor (2010)	2010	Global	Daily demand for pirated e-books	1.5-3 million people
Live Football Matches	Net Result (2011)	2010-2011	Global	Average number of websites providing infringing content	197.1

Notes: (*): information not available in the source document.

There is a major distinction between methodologies that aim to estimate the size of counterfeiting and piracy, and methodologies that aim to estimate their effects. The outputs of the former are usually inputs for the latter. Some of the sources we reviewed are not very clear in separating these two aspects, mainly because some of the metrics that capture these phenomena are used interchangeably. The frontier between these two aspects is represented by lost revenues. It can be considered as a measure of the size of counterfeiting and piracy, but it is also a first-order effect of it. There is also a clear imbalance in that many studies draw on already existing estimates of the size in order to say something about the effects, where the latter seems a less demanding process than estimating the size itself.

A number of studies that attempt to measure the effects of counterfeiting and piracy were critiqued in the 2008 OECD report on counterfeiting, so in what follows we will mainly focus on aspects that were not considered in that report and on studies that were released later. Table 3-2 summarises some of the potential types and estimates of impacts that have been identified in the literature. This table should not be considered as an exhaustive overview of potential effects or their estimates, rather it is indicative of the variety of impacts, of the metrics used and the range of estimates that have been produced.

Also, we do not consider the vast literature that analyses the effects of the strength of IPR protection. As the 2008 OECD study shows, the number of IPR infringements and the strength of IPR protection regimes certainly seem to be correlated, but not as strongly as one would expect (OECD, 2008 p.108). Therefore, it does not appear correct to treat estimates of the effects of the strength of IPR protection regimes as estimates of the effects of counterfeiting and piracy, because the two dimensions are not perfectly correlated.

3.3.1 Existing research on the effects of counterfeiting

Two broad types of counterfeit products have been identified (Wall and Large, 2010):

- Safety-critical goods – including medicines, fertilisers, aircraft and car parts, as well as cigarettes, alcoholic beverages, food, soft drinks, mineral water, toys, cosmetics and sunglasses

- Counterfeiting of luxury goods – including clothes, bags, shoes, watches, jewellery, leather goods and so forth.

Although comparatively little has been written about this, safety-critical counterfeits constitute an important challenge for governments and law enforcement precisely because they pose important and immediate public health risks. Consumers may suffer injury or death from poor quality counterfeit goods. For instance, the Pharmaceutical Security Institute, a membership organization working in the area of counterfeit medicines, has argued that the current and future risks from counterfeit medicines include: treatment failure in malaria, tuberculosis and HIV/AIDS; growth of resistance to existing anti-infective medicines from the use of sub-par treatments; and the spread of drug-resistant pandemics, such as HIV and influenza (Pharmaceutical Security Institute, 2005). However, as an OECD report argues, 'because data are not being collected systematically, most evidence of negative health and safety effects is anecdotal in character and more work is needed to measure the effects more broadly' (OECD, 2008).

Table 3-2:   Examples of types and estimates of impacts identified in the literature

Type of impact	Source	Estimate	Countries	Additional information
Macro effects				
Economic growth	CEBR (2002)	€8bn decrease in EU GDP	EU	Reduction in investment in innovation due to counterfeiting (and piracy).
Foreign direct investment	OECD (2008)	n/a[6]	Germany, Japan, US	ATRIC regression analysis for foreign direct investment originating in Germany, Japan and the US indicates that higher rates of investment from these countries are correlated with lower rates of counterfeiting and piracy.
Output (revenue and related measures of economic performance)	IPI (2007)	$12.5bn per annum	US	Due to IPR infringements in the sound recording industry.
Output	Siwek (2006)	$20.5bn per annum	US	Due to IPR infringements in the motion picture and related industries.
Public finance effects				
Customs revenues	OECD (2008)	n/a	Global	
Tax revenues and higher welfare spending	Frontier Economics (2011)	$77.5bn per annum	G20	
Tax revenues	Frontier Economics (2011)	$125m per annum	G20	
Tax revenues	IPI (2007)	$422m per annum	US	Due to IPR infringements in the sound recording industry, comprises personal income, corporate income and production taxes.
Tax revenues	Siwek (2006)	$837m per annum	US	Due to piracy in the motion picture industry, comprises personal income, corporate income, production and sales taxes.
Tax Revenues	Thompson (2004)	$2.6bn in 2003	New York State and New York City	
Tax revenues	US Chamber of Commerce (2006)	BRL12bn per annum for the apparel and toys industries	Brazil	

---

[6] The point estimate has no straightforward interpretation because the main independent variable is a ranking

Type of impact	Source	Estimate	Countries	Additional information
Increased costs of prevention and combating of crime	Frontier Economics (2011)	$25bn per annum	G20	
Costs associated with efforts to control counterfeiting	Frontier Economics (2011)	More than €20bn for every 1% increase in the crime rate caused by the trade in counterfeit and pirated goods	G20	Including indirect costs related to: judicial process, investigation and enforcement; raising of public awareness and the management of confiscation both of illicit goods and their means of production.
Producers and consumers: cost-benefit effects				
Revenues of IPR holders	International Trademark Association (1998)	22% of total sales	Global	Estimate based on 1995 data, as average lost sales as % of total sales in the apparel and footwear industry
Revenues of IPR holders	Rob and Waldfogel (2006)	an additional download reduces sales by between .1 and .2 units	US	Music industry. Based on a convenience sample of college students. The exact quantification of the net effect of illegal downloads is debated.
Revenues of IPR holders	Informa Media Group (2005)	$2.1bn per annum	US	Music industry, due to illegal file-sharing.
Revenues of IPR holders	LEK Consulting as reported in Motion Picture Association of America (2006)	$ 6.1 bln per annum	Global	Estimate based on 2005 data for the motion picture industry
Revenues of IPR holders	Jupiter Research (in IFPI, 2009)	£180m in 2008	UK	Film, TV and music industries.
Revenues of IPR holders	Centre for Economics and Business Research (2000)	Euro 1,554 million per annum	EU	Estimate based on 1998 data for the pharmaceuticals industry
Revenues of IPR holders	Centre for Economics and Business Research (2000)	Euro 3,017 million per annum	EU	Estimate based on 1998 data for the perfumes and cosmetics industry
Revenues of IPR holders	Centre for Economics and Business Research (2000)	Euro 3,731 million per annum	EU	Estimate based on 1998 data for the toys and sports equipment industry
Revenues of IPR holders	Centre for Economics and Business Research (2000)	Euro 7,581 million per annum	EU	Estimate based on 1998 data for the apparel and footwear industry

Type of impact	Source	Estimate	Countries	Additional information
Revenues of IPR holders	Tera Consulting (2010)	€10bn per annum	EU	EU creative industries.
Costs of removing malicious software due to piracy	BSA/IDC (2010), IFPI (2004)	n/a	-	Including violations of privacy and damage to the computer itself.
Population health effects				
Population health	Satchwell (2004)	n/a	Global	Diversion of medicines that were directed to poor countries; higher number of malaria deaths due to ineffective drugs; gains from the trade of counterfeit medicines used to fund other crimes; growing number of hospital admissions due to adverse reaction to medication.
Consumer health and safety	OECD (2008), Forzley (2003)	n/a	Global	Health risks posed by counterfeiting, especially automotive, food, pharmaceutical.
Global cigarette consumption	Joossens (2009)	2% higher than without counterfeiting	Global	Presuming the effect of counterfeiting on cigarette prices.
Deaths from cigarette consumption	Joossens (2009), based on estimates from Mathers and Loncar (2006)	Up to 164,000 averted deaths yearly from 2030 onwards	Global	
Deaths resulting from counterfeiting	Frontier Economics (2011)	$18.1bn per annum	G20	Economic cost estimate of deaths resulting from counterfeiting.
Additional cost of health services caused by dangerous counterfeit products	Frontier Economics (2011)	$125m per annum	G20	
Impact on occupational safety	Wall and Large (2010)	n/a	Global	The production of counterfeited goods may occur in poor working conditions, with employers not abiding by a country's labour standards.

Type of impact	Source	Estimate	Countries	Additional information
		Employment effects		
Employment	CEBR (2002)	17,000 jobs across the following sectors: clothing and footwear, perfumes and toiletries, toys and sports goods, and pharmaceuticals	EU	On the other hand, in low-income countries with little IPR, the manufacture of counterfeited and pirated products may in fact generate employment and income, though working conditions may be unstable and/or unsafe.
Employment	Siwek (2006)	Up to 120,085 jobs in the motion picture industry and up to an additional 20,945 in related industries	US	Due to piracy in the motion picture industry.
Employment	Siwek (2006)	Up to 71,060 jobs in the sound recording and related industries	US	Due to IPR infringements in the sound recording and related industries.
Employment	Frontier (2011)	2.5m jobs across the economies	G20	Due to counterfeiting and piracy.
Employment	TERA Consulting (2010)	185,000 jobs	EU	Due to counterfeiting and piracy of film, TV series, recorded music and software.
Employment	Europe Economics (2008)	30,000 jobs	UK	Due to music, film and TV piracy.

According to Wall and Large (2010), counterfeit luxury products do not present equally urgent public health dangers, although it is clear that they may not adhere to the same quality standards as the original. However, the production of these goods may occur in poor working conditions, with employers not abiding by a country's labour standards and exposing their workers to various types of occupational risks (such as exposure to toxic products, injuries with machinery, etc.), although these risks are of course not exclusive to the production of counterfeit products.

Counterfeit goods of both types pose other risks to consumers, most notably the risk of paying high sums for low quality products which do not perform as expected. However, as Wall and Large (2010) argue, specifically with reference to luxury goods, that not all counterfeits are of equal quality and level of deception to consumers, nor do they all pose the same risk to consumers. The authors argue that the highest risk to consumers is presented by counterfeit goods that are of low quality but highly deceptive; the potential buyer purchases a good she believes to be legitimate, but which is of much lower quality than the original would be. This places consumers at financial risk because 'the goods are often still very expensive, but it also creates bad feelings towards the brand owner when the merchandise fails to live up to expectations' (Ibid., p.1106). This is unlike low deception goods, where the customer is almost always aware that she is buying a counterfeit product. High deception counterfeit goods are also 'most representative of lost sales and are therefore the most logical group upon which to base estimated of financial losses' to counterfeiting (Ibid., p.1107).

Counterfeit goods may also have an impact on innovation and growth, although the extent of this impact remains unclear. Illicit businesses draw on legitimate producers' investments in technology, product development, marketing and brand value. All of these activities are time-consuming and investment-intensive, and do not always bring immediate results. This is why new ideas and products are protected by patents, copyrights, design rights and trademarks, so innovators can fully exploit the eventual benefit of their long-term investment costs. Counterfeiting poses a challenge to this. For instance, one estimate by the Centre for Economics and Business Research (CEBR) suggests that counterfeiting (and piracy) can lead to a reduction in investment in innovation, which could in turn have a negative effect on GDP across the EU of €8 billion (CEBR, 2002).

The possible effects of counterfeiting on employment are also worth considering. The CEBR study quoted above, for example, estimates that the impact of counterfeiting on employment in the EU translates into the loss of 17,000 jobs across four sectors: clothing and footwear, perfumes and toiletries, toys and sports goods, and pharmaceuticals. The impact on employment, however, may not be equally distributed at the global level. In third world countries with an abundance of cheap labour but little home-grown intellectual property, the production and distribution of counterfeit and pirated products may in fact generate employment and income, though working conditions may be unstable and/or unsafe.

The production, distribution and sale of counterfeit products also have complex and potentially significant direct impacts on the finances of governments. The most important of these include forgone tax revenues and the additional costs incurred by law enforcement and justice systems when combating counterfeiting operations. Forgone tax revenues occur in a variety of ways. Illicit products are usually manufactured using undeclared work in the black economy. Furthermore, counterfeiters often target products that attract high rates of taxation, such as tobacco and alcohol. To the extent that these may reduce sales volumes for legitimate goods, this reduces the

tax revenue from the sale of the legitimate goods. Counterfeit goods smuggled across borders involve a loss of custom revenues (OECD, 2008). A Frontier Economics study that expanded on the OECD's 2008 report estimated that counterfeiting and piracy may cost G20 governments and costumers over $120 billion every year, with $77.5 billion of this from tax revenue losses, $25 billion in increased costs of crime, $18.1 billion in the economic cost of deaths resulting from counterfeiting and $125 million for the additional cost of health services caused by dangerous counterfeit products (Frontier Economics, 2011). Studies have also estimated the impact of counterfeiting and piracy on tax revenues in more limited geographical areas. A 2006 study conducted by the US Chamber of Commerce estimated that forgone tax revenues in Brazil amounted to BRL12 billion (= US$5.9 billion at current rates) for the toys and apparel industries (US Chamber of Commerce, 2006). One study estimated the loss of tax revenues in New York City and New York State at $2.3 billion (Thompson, 2004).

Finally, governments also incur a range of additional costs associated with efforts to control counterfeiting: these may take the form of judicial process, investigation and enforcement, raising public awareness and the management of confiscation both of illicit goods and their means of production, and so forth. In the G20 countries, the economic and social costs of crime caused by the trade in counterfeit and pirated goods were estimated to increase by more than €20 billion for each 1% increase in the crime rate (Frontier Economics, 2011).

3.3.2 Existing research on the effects of UUPC

Unauthorised use of content such as music, films, software or broadcasting generally does not tend to be safety-critical: it does not present health risks to consumers. In addition, the limited requirements to manufacturing and distribution for such pirated goods, particularly for online piracy, mean that labour-related issues (such as workers' conditions) are not a major concern when it comes to this IPR violation. Rather, the main concern with piracy tends to be, primarily, the financial risk that it poses to IPR holders in the form of lost revenue. For instance, recording industry revenue fell sharply in the last decade, which many (but not all) analysts attribute to extensive illegal downloading of music files over the Internet (Rob and Waldfogel, 2006). Research literature, however, is yet to paint an updated picture of how these industries have adapted and innovated in the last few years by co-opting file-sharing to create new revenue streams.

There are some similar costs to those identified in the case of counterfeiting, however. For example, one study examining film piracy in the US estimates that this activity resulted in a direct loss of over 120,000 jobs in the US motion picture industry, with additional losses in other affected sectors (Siwek, 2007). Thus lost legitimate employment, lost tax revenue and increased government spending to control and deter piracy are all economic consequences of this activity. In the case of software piracy, however, consumers also face risks. When software download from P2P networks is accompanied by attempts to install malicious or unwanted software in the consumer's computer, it can take time and money to recover any lost data and information. If such an incident occurs within an organisation, the loss or corruption of data and information can cost significantly more, and potentially damage the company's reputation.

Unlike the literature on counterfeiting, research on music, film and software piracy has been, to a certain extent, fraught with debate over the true impact of this IPR violation. There is considerable debate over the question of whether piracy may actually incur some positive externalities for business. These would occur, for instance, as pirated products enable individuals

to 'sample' a product and then purchase the legitimate version, or allow certain people to adopt or consume a new product at a cheaper price, thus 'spreading the word' about it and increasing the popularity and demand for its legitimate version (Hui and Png, 2003; Oberholzer-Gee and Strumpf, 2004). In the case of software piracy, for instance, network externalities may occur, whereby consumers derive additional utility from an increased user-base for the software (Haruvy *et al.,* 2004). This is because a larger user network for particular software enables file compatibility and transferability, reduced uncertainty about product quality, a larger base of compatible software, and so forth (ibid.). For illegal music downloads, the balance between the utility gains for consumers and the losses to the recording industry depends at least in part on which albums are downloaded. As Rob and Waldfogel (2006) have stated:

> If downloading tends to occur for albums that consumers value highly and would otherwise have purchased, then revenues are reduced … However, if downloading tends to occur for low-valued albums, downloaded albums are not candidates for being purchased in the first place … .

The authors' empirical research suggests that while piracy does displace sales, this displacement is incomplete (i.e. not a one-to-one relationship between illegal downloads and purchased music) in part because downloaded music is less valued than purchased music (ibid.). The balance of the positive (increase in legitimate demand) and negative (lost revenue) effects of piracy in all of these industries remains an empirical question for further research.

3.4   In sum

Based on the literature reviewed, this chapter has set out findings from existing research on drivers of supply of and demand for counterfeited and pirated products, noting that the significant and rapid technological advances and economic growth in emerging countries of recent years have most likely facilitated, sped up and widened the scale and scope of counterfeiting and UUPC globally, presenting fresh and difficult challenges for policy and practice. Key macro-economic supply drivers that have been identified in the literature include the growing prevalence of digital and networked technologies, the globalisation of trade, the growing importance of international brands, the presence of large integrated markets supporting free trade, low or weak enforcement of penalties targeting violators of IPRs and the growing presence/involvement of organised crime in the production and distribution of counterfeited and pirated goods. In addition to these macro-economic factors additional industry-specific factors have also been identified. Supply factors are not the only ones influencing the production and sale of counterfeited products and UUPC, however.

Demand for counterfeit products and the prevalence of UUPC are also relevant, and research has shown that there are a number of relevant macro-level factors as well, including a population's willingness to buy counterfeit goods, the lack of available authentic goods, the high price of authentic goods and the rising quality of counterfeit goods.

The impacts of counterfeit and UUPC could have important implications for consumers, safety, policy and decision-making. A necessary condition for studying impacts is having consistent estimates of the size of counterfeiting and UUPC themselves, an issue still lacking the necessary scientific consensus as discussed throughout the rest of this report. In our reviews of how such estimates could be used to infer impacts, different paths and concerns have emerged:

- Econometric models that rely on regression analyses, such as the ones used by the OECD (2008), are relatively easy to implement but need care in the interpretation of results and sufficient geographical and chronological coverage to achieve statistical precision.

- Macro-economic simulation models provide an avenue for the estimation of economy-wide impacts, such as those on employment and fiscal revenues. However, they have often lacked the necessary transparency (e.g. the 2002 CEBR study mentioned above, as well as The Allen Consulting Group (2003) and KPMG (2008) reviewed in detail in Appendix A). Or, where their methodologies have been rigorously described, studies have nevertheless been criticised for some of their assumptions (as in the critique of the Siwek study by the US Government Accountability Office (2010)).

CHAPTER 4 What existing methods can teach us

In order to understand better existing approaches to estimating the scale, scope and impacts of IPR infringements, we conducted an extensive review of literature that examined the different methodologies which have been used for deriving these estimates. By 'scope', we mean the array of product categories subject to IPR infringements, as well as the geographical spectrum of the locations of producers and consumers of counterfeit and pirated goods. Under 'size/scale' we include the actual volumes of infringing goods and/or monetary measures of lost sales. By 'impacts', we refer to the variety of consequences that have been either theoretically or quantitatively related to counterfeiting and UUPC. IPR violations and their different types have been identified and defined in Chapter 2.

Findings from the review presented here set the stage for the improvement of existing approaches or the development of new ones. This chapter provides a brief overview of a few common methodological and data issues affecting estimates of both counterfeiting and UUPC, presents detailed reviews of the methodological approaches used to date, and synthesises key findings. Appendix A provides a detailed assessment of individual literature sources as well as an annotated bibliography.

4.1 Methodology of our literature review

The individual papers we review in this section and elsewhere in this report were retrieved from an extensive desk-based search for literature, using the subscription and non-subscription databases in the RAND Library (an online library with access to over 100 subscription databases and over 1,000 individual journals). Key search terms deployed in various combinations, included: *intellectual property right, copyright, methodology, piracy* (the term used in the literature on UUPC), *counterfeit*, estimate, magnitude, size, scale, scope, impact, effect* and *research*. Additional documents were obtained from members of the European Observatory on Counterfeiting and Piracy, who sent relevant peer-reviewed papers, grey literature, industry reports, presentations and media articles to the research team. The papers included in this section here are limited to those which provide original attempts to estimate quantitatively the scope, size or effects of counterfeiting and/or piracy, in any geographical area, and for one or more products.

The literature search resulted in nearly 250 publications, 80 of which were analysed in detail as they matched the criteria outlined above. The review of each paper consisted of the extraction of key information that was recorded for each document, which included all bibliographical information, type of publication, research question (i.e. focus of the estimation), methodological approach, data used, geographical and product focus, and year/s of analysis. We then assessed

each paper on the basis of the robustness and suitability of the methodological approach to address the paper's research question, taking into account the methodology as described in the paper, the data used and other considerations.

While our literature searches were extensive, it is possible that relevant papers were missed. Reasons for this would include papers published in sources to which the RAND Library does not have access, papers published in languages other than English and papers published after we conducted our search. In spite of this limitation, this review provides an informative overview of the types of methodologies that have been used so far, and is thus a useful starting point for the development of our own methodology to estimate the scope of IPR violations.

4.2   Assumptions in estimations of the scale of IPR violation

Estimates of the scale of IPR violation generally use demand-side or supply-side approaches. Demand-side approaches are typically based on consumer surveys that try to elicit preferences and actual behaviours. While this type of approach has the benefit of identifying some counterfeit or UUPC exchanges that go unnoticed or unrecognised by law enforcement, its accuracy is compromised by under-reporting, misreporting, and lack of respondents' knowledge of the frequency of occurrences (for purchasing counterfeit products). These methodologies are discussed in detail in a study by the Centre for Economics and Business Research (CEBR, 2002). On the other hand, supply-side approaches, which try to quantify the volume of counterfeits available to the market by a measure of trade, typically seizures (OECD, 2008) or some parameter on the incidence of counterfeiting in a given industry, have their own limitations. In particular, most of these analyses start with some basic measure of known counterfeiting (either through a seizure rate or hypothesis), which will never fully reflect the total size of the market.

There is no way of observing either demand or supply of counterfeits directly, so no best approach exists. Extrapolations for missing or under-reported information are necessary using either approach, and the uncertainty of assumptions on which these extrapolations are based remains an issue.

There are very few studies that attempt to produce an original estimate of the magnitude of counterfeiting and UUPC. The OECD's (2008) is the only one to attempt a comprehensive estimation, covering a large sample of countries and all sectors subject to counterfeiting (considering that the Frontier Economics (2011) study just expands on it and does not present an alternative methodology, while the CEBR (2002) study is a 'design' exercise, not an estimation one). This speaks to the intrinsic difficulty in such estimation exercises.

4.3   Data issues in estimating the scale of counterfeiting and UUPC

The existence and reliability of data is one of the most pressing problems for estimating the magnitude of counterfeiting and UUPC. Nevertheless, there are a number of different data sources from which information on the magnitude can be extracted. The World Intellectual Property Organization (WIPO) and OECD (Olsen, 2005) list the following:

- enforcement information;

- surveys of consumers, producers, and/or distributors;

- sampling and testing of products;

- economic models.

As with any estimation, each approach and data source has its limitations. These are described below.

4.3.1 Enforcement information

Enforcement information includes information, not just on the seizure of counterfeit or pirated goods, but also on legal actions pursued by legitimate producers. Such data are valuable for providing a basis for understanding the range of products that might be impacted by counterfeiting and UUPC, but do not provide a good measure of the magnitude of the problem. This is because enforcement data reflect an unknown fraction of the total amount of counterfeiting and UUPC that actually occurs, and also reflect the subset of cases in which people/organizations are caught. Nonetheless, such data provide valuable insights as to the range of products traded and the geographical regions in which counterfeiting and UUPC are occurring.

An example of this methodology is illustrated in the GAO (2010) report, which assesses the scope of counterfeiting and piracy through the Customs and Border Protection data. In another example, the OECD (2008) customs survey asked officials to report seizures. One important issue in the OECD survey was that the customs officials in only 13 countries were able to report data on seizures using the internationally recognized standard, the six-digit Harmonized System developed by the World Customs Organization; this represents an important limitation to the comprehensiveness of cross-country comparisons. However, as with the GAO report, the OECD study provides important insights into the types of products that are counterfeited and pirated, as well as exporter countries or regions.

4.3.2 Surveys of consumers

Consumer surveys are well-suited for illuminating consumption and purchase patterns of particular goods, consumers' attitudes towards fake products and their willingness to engage in illegal activity to obtain pirated or counterfeit products for cheaper prices than the genuine item. Examples of these efforts can be found in the academic literature, which highlights the importance of non-economic factors. Cordell *et al.* (1996) surveyed 221 business students and found that willingness to purchase counterfeits is negatively related to attitudes towards lawfulness. Poddar *et al.* (2011) show that not only price motives but also moral beliefs matter for the decision to purchase counterfeits, and in particular the consumer's perceptions of the right holder's corporate responsibility.

Aside from statistical considerations which, in the absence of a representative sample, undermine the external validity of any estimate, the intrinsic limitation of survey data stems from the fact that respondents are unlikely truthfully to report purchases of counterfeit or pirated goods. Randomised response design techniques can help overcome some of these concerns. A second weakness is that consumers cannot provide useful information on the extent to which they unknowingly purchased fake products. In fact, a recent independent review commissioned by the UK government that focuses on online piracy in the UK, states that 'in four months of evidence gathering, we have failed to find a single UK survey that is demonstrably statistically robust'; moreover, '[f]or many surveys, methodology is not available for peer review' (Hargreaves, 2011,

p.69). Tellingly, the report concludes that 'we have not found either a figure for the prevalence and impact of piracy worldwide or for the UK in which we can place our confidence' (*ibid,* p.73).

Many studies limit the scope of surveys by exogenously restricting their focus to a limited number of product categories, such as luxury consumer goods, software or pharmaceutical products. These choices are legitimately dictated by the interest of the researcher or funding source, or by convenience. The Gallup survey (Gallup Organization, 2007) represents an exception in that it provides two important insights on the scope of counterfeiting and piracy.[7] First, the study finds that globally, the main three categories of counterfeit and pirated products are: 1) fashion, bags and footwear; 2) music; and 3) movies. Second, it finds that there are differences in preferences across countries for the types of counterfeit and pirated goods.

At best, if individuals do not under-report or misreport, surveys of this kind allow us to draw inferences about mean probability of purchase of a counterfeit good in the population and mean frequency. However, the limitations of these types of survey make them more reliable as a measure of the *variation* in the geographical and sectoral scope of counterfeiting and UUPC, rather than of the exact quantities. They also offer valuable information on consumer attitudes to and degree of awareness of the phenomena, as well as on distribution channels.

4.3.3 Surveys of producers and distributors

Producer/distributor surveys are useful in markets where consumer awareness of deception is low and unusual and where non-market-driven disruptions impact sales of particular products. Indeed if producers notice declines in sales that are outside the range predicted by standard market models, it may be indicative of a counterfeiting and UUPC problem. However, counterfeiting and UUPC are likely to be noticed only when infringement on market space is fairly high and done at a time when other market forces are fairly stable.

One example of a study using this type of survey is the Swiss Federal Institute of Intellectual Property (2004) study, which surveys businesses about the nature and extent of counterfeiting and piracy from which they suffer both domestically and internationally and thus could in principle lead to estimating sector-specific incidence factors for counterfeiting and UUPC. However, the study has a very limited sample size (72) which seems to lend no power to any category-specific inference. The OECD (2008) study also uses and draws on industry surveys to understand in which sectors IPR violations occur, thus shedding additional light on the scope of the problem. However, the sample sizes of these surveys are not disclosed.

4.3.4 Sampling

Also known as mystery shopping, this approach involves having consumers go out and buy products from various vendors. It is an effective way of learning first-hand the extent to which IPR infringements might occur, since with this method the researcher knows the sampling method and therefore can infer something about the scale of the problem. Indeed, they are especially productive for items for which deception of the consumer is high. However, it is a very costly approach, particularly if multiple markets are being considered simultaneously. An example of this methodology is the European Alliance for Access to Safe Medicines (2008) study, which investigated online pharmaceutical sellers.

---

[7] It derives the scope from the answers of the respondents instead of restricting the scope of the questions beforehand.

4.3.5   Economic models

Economic models to predict counterfeiting and UUPC use data obtained from the sources mentioned above and build mathematical models following the logic of economics, but incorporating elements of uncertainty, to predict what is unknown about the specific market under scrutiny. The sophistication of these models varies substantially. For example, they may be as simple as subtracting legitimate sales from an estimate of total demand (predicted from a model of demand). However, where there are reliable data on demand for multiple products susceptible to IPR infringements in certain countries, econometric country-level models can be estimated and then potentially used to extrapolate for countries that lack data (Qian, 2011). While specific sectors have been studied using these approaches, no broad formulation of these models has been considered for estimating counterfeiting and UUPC across products, sectors and markets.

The key advantage of estimates based on economic models is that they can incorporate a variety of factors considered simultaneously (such as unknown quality, price, replacement ratio) to project anticipated sizes of the market for either different goods or a country as a whole. Further, they can be easily modified, and at relatively low cost, to consider alternative factors not yet incorporated in the model. The disadvantage is that these models are only as good as the logic and data that underpin them, and suffer from whatever omitted variable biases might exist for factors not considered or those that are poorly measured. Moreover, they are predictive, not actual estimates. Hence, depending on the size of the error band around the prediction, the model may be more or less informative for policymaking.

Examples of papers that model the extent of IPR infringements include Shapiro and Varian (1999), who model file-sharing as a form of free-sampling, which may stimulate sales; Conner and Rumelt (1991), who model utility as a function of the size of the user base, which results in positive demand effects of piracy; and Raustiala and Sprigman (2009), who theorise about the 'piracy paradox' in the fashion industry, where much innovation occurs in the presence of high levels of copying of design items. For a thorough review of papers modelling IPR infringements, please see Pacula *et al.* (2012).

4.4   Methodological issues regarding estimates of the effects of IPR infringements

The OECD (2008) study mentioned earlier, as well as developing a methodology for estimating the size of counterfeiting and piracy, attempts to explore their effects on foreign direct investment (FDI) and trade. An index of counterfeiting and piracy is computed for specific sectors, based on seizure data collected by customs in an analogous fashion to the index used to estimate the size of counterfeiting and piracy.[8] Ordinary least squares regressions are then run: the dependent variable is the share of products of a given type in the total exports of a given economy, and the main independent variable is the index of counterfeiting and piracy. The analyses only found

---

[8] This index is called ATRIC (aggregated trade-related index of counterfeiting and piracy) and is somewhat different from G-TRIC (general trade-related index of counterfeiting and piracy). One important difference is that the former is computed at the economy-by-sector level, while the latter is computed separately at the economy and at the sector level, in order to maximize coverage. As the main focus of this section is on the methodologies to estimate the effects, we refer the reader to the 2008 OECD report for the details on the construction of the ATRIC.

significant negative effects of counterfeiting and piracy on the shares of pharmaceutical products.[9] The results for very few other product categories are reported, and in these regressions the coefficients of the index of counterfeiting and piracy are only marginally significant. This suggests that the evidence of the impacts of counterfeiting and piracy (as measured in this study) on trade is limited. Linear regressions are also employed for the analysis of the effects of counterfeiting and piracy on incoming FDI. The level of incoming FDI from Germany, Japan and the US is regressed on the index of counterfeiting and piracy, controlling for population size and a measure of openness (GDP per capita to total trade). The results suggest a negative impact of counterfeiting and piracy on the inflow of FDI from Germany, Japan and the US, and are robust over a few different specifications.

However, these regressions suffer from two important limitations, namely the small sample sizes and the ambiguity in the interpretation of the results. While the former is an intrinsic limitation that any study of IPR infringements will encounter, due to the lack of data, the latter arises from the difficulty of isolating the different mechanisms at play that might relate counterfeiting and piracy to trade and FDI. Particular care should be taken in explaining the potential effects related to omitted variables, as well as to using a language that avoids implying anything stronger than an association conditional on other variables. For example, the wording in the OECD report -– 'a higher rank for counterfeiting translates into smaller FDI flows' – (OECD, 2008 p.167) is misleading, because it suggests that the analysis captures some sort of a transmission (cause-effect) mechanism.

Overall, one of the main weaknesses of some of the existing estimates of the effects of counterfeiting and piracy on macro-economic variables such as GDP and employment is the assumption of a 100 percent substitution rate between counterfeit and genuine products – see for example the OECD (2008) critique of a 2005 IDC study. In addition, one aspect that seems to be systematically excluded from existing studies is consumer surplus. Consumer surplus refers to the welfare benefit of getting access to a substitute good at a lower price, while many studies consider the negative effects of counterfeiting and piracy on consumers (unemployment and health and safety risks) and on producers (lost revenues). Huygen *et al.* (2009) provide an example of a comprehensive treatment of the distribution of welfare effects and of their net balance (i.e., the balance between costs and benefits) in the case of file sharing (Huygen *et al.,* 2009). From an economics point of view, any study that neglects consumer surplus in a welfare analysis is incomplete.

A CEBR (2000) study develops a model to assess the impact of counterfeiting on GDP and employment. It models the impact of counterfeiting as a negative investment shock over five years equal to the reduction in investment that achieves the average rate of return on capital given the reduction in profits estimated earlier in the study. This negative investment shock is plugged into a macro-simulation model that computes the effects on employment and GDP. The results, described more completely in the previous section, suggest a loss of 17,000 jobs in four sectors. The model is based on the national accounting framework and can be populated with data from the European System of Accounts.

The Allen Consulting Group (2003) study built a multi-sector dynamic macro model. A reduction of 33 percent in counterfeiting and piracy over five years was studied as a deviation

---

[9] Significant at 10 percent significance level.

from the base-case scenario. Two mechanisms were triggered. The first was a negative macro-effect of the expenditure shift towards genuine products: the effect is negative because in the long run the economy relies more on imports (for example of genuine software if counterfeiting is reduced) and devaluation of the exchange rate, thus increasing exports, ultimately reducing the price of exports and thus worsening terms of trade, which has recessive effects on the economy. The second is a positive investment shock (we assume due to the reduction in counterfeiting and piracy which increases incentives to research and development, but this is not explicitly explained), which also boosts employment.

The 2008 KPMG study also incorporates a model of the economy. In particular, it developed a macro-econometric model to study effects of counterfeit goods on the overall economy. A number of macro-economic variables are modelled (GDP, retail trade, tax collections, employment and household expenditure) as a system of simultaneous equations. The effect of eliminating counterfeiting on these variables is introduced as an increase in the demand for genuine goods under the assumption that the substitution rate is 1. A 10 percent figure is assumed as the overall level of counterfeiting. The same exercise is repeated at the sector level.

All these three studies do not present the details of their respective macro-economic models. This lack of transparency regarding the approach makes it impossible to review and comment on their quality and suitability.

Surveys have also been used to gain insights into the relative, rather than absolute, magnitudes of the effects of counterfeiting and piracy. For instance, OECD surveys of governments and industries highlight the importance of reduced profitability and hazards to health and safety (OECD, 2008). It is important to note, however, that on their own surveys of this kind may provide information on the types of risks or effects of counterfeiting and UUPC, but are unlikely to be able to provide reliable estimates of the scale of these effects.

Methodological concerns also arise from estimates of sector-specific effects of counterfeiting and piracy. A methodology to estimate the effects of counterfeiting and piracy that relies on macro-economic modelling is proposed by Siwek in the 2007 Institute for Police Innovation study on sound recording. Siwek applies multipliers from the RIMS II mathematical model of the US Bureau of Economic Analysis.[10] In particular, he selects multipliers to look at the effects on output, earnings and employment on all US sectors, as well as at the direct effects on earnings and employment in the sound recording sector. The inputs to this methodology, apart from the multipliers themselves which are taken from the model, are estimates of the monetary losses from piracy in the sound recording industry (see 'Institute for Policy Innovation (IPI) (2007)' in Appendix A). The author also estimates tax losses from personal income taxes that the sound recording industry employees would have paid and from corporate income taxes. Lost personal income taxes are computed by applying an average tax rate (total personal current taxes paid in the US over total US personal income) to the lost employee earnings. A similar operation is done for corporate income taxes. The GAO (2010) report critiques the use of multipliers for these purposes because it ignores both the fact that resources in expanding industries must be drawn from other sectors in the economy and the effects of changes in consumer income on demand.

---

[10] The Regional Input–Output Modeling System is a method to estimate regional input–output multipliers. The multipliers are used to study the economic impact of changes in output, earnings or employment in a given sector on total output, earnings and employment in the economy, accounting for the economic relationships between sectors.

Still, this is by far the most comprehensive attempt found in the literature to quantify the effects of counterfeiting and UUPC.

4.5   Synthesis of review findings

This section presents the key messages emerging from our review of methodologies used in existing research to estimate the size, scope and impact of counterfeiting and UUPC.

Proxy indicators are needed to assess the magnitude of illicit markets
In illicit markets, it is not possible to observe either demand or supply of counterfeits and pirated products directly. This means that proxy indicators, i.e. indirect measures that approximate or represent a phenomenon in the absence of direct measures, will be required to estimate production, trade and consumption in these markets. Extrapolation methods are then required to estimate volumes of IPR infringements. Moreover, as data are often missing, incomplete or inconsistent over time or across geographic regions, it will be necessary to triangulate information from alternative approaches and sources. However, not all proxy indicators are equally suitable or able to provide us with reliable and accurate estimates.

At present, data sources in counterfeiting and UUPC are primarily based on consumer surveys and seizures
Most attempts to estimate the size of the counterfeit and UUPC markets draw largely on self-reported information from consumers and suppliers as well as from law enforcement data on known seizures and law suits. More sophisticated efforts attempt to merge these data sources and/or use them in conjunction with information obtained through mystery shoppers (in the case of counterfeiting) or sting operations (in the case of UUPC). All of the studies must, to some extent, rely on mathematical assumptions to deal with under-reporting or missing data either within a product-line, within an industry, or across countries. Whether these assumptions are accurate obviously varies by the quality of information that is known versus that which is unknown. In future, however, new estimation methodologies might be developed that rely on different data sources (for example, Pacula *et al.,* 2012).

There is little convergence on methodology in counterfeiting literature
It is clear from the review of the counterfeiting literature that, while numerous attempts to approximate the scale of the problem have been made, there has been little convergence on a preferred methodology and innovation in methods and forms of collecting data continues. Furthermore it is often difficult to assess the quality of specific studies in this literature, as there is frequently a lack of transparency about assumptions, data values or sources.

There seems to be more convergence on methodology in the literature on copyright infringements
There appears to be more convergence on methodologies for estimating the extent of unauthorised use of online content than is the case with counterfeiting. The reason for this may be that 'piracy' today is dominated by online copyright infringements, predominantly illegal downloading or file-sharing. This has fundamental consequences for the distribution channels of pirated products. Therefore, in contrast to counterfeiting and physical pirated products such as CDs and DVDs, using seizures as a proxy for trade is not suitable for estimations of the magnitude of online UUPC. Estimates for these copyright infringements typically utilise data on the number of legal products sold, the UUPC rate (obtained through survey methods or law

enforcement investigations) and the number of actual downloads. Generally speaking, this appears to be a sensible approach when seeking to understand the number of 'pirated' goods across a range of product types (for example, music, film and broadcasts) produced around the world and consumed around the world. However, the nature of how this information is then used to provide an accurate estimate may be an area for further exploration. In particular, there is scope for considering a slightly innovative approach to bring about a more reliable range for the estimate. However, similar to the literature on counterfeiting, there is lack of clarity and consistency in how extrapolations to specific markets or countries are performed, in large part because the literature on the determinants of UUPC has not yet reached consensus on what drives UUPC. For example, the research analysed thus far provides divergent findings as to whether broadband penetration and Internet access increase or decrease UUPC rates; there is some indication that these factors must be considered in conjunction with the speed of economic development. Furthermore, if the factors were clearly identified, then researchers could use and build off the approach in order to provide estimates in other areas or more accurate estimates.

Understanding knowledge level by consumers is crucial to elucidate substitution rates
The assumption regarding the extent to which substitution occurs within a counterfeit market is an area where much work remains to be done. It is perhaps best done considering goods on a product-by-product basis, as observed quality often varies by type of product. Furthermore, it is critical to have knowledge about consumer 'types' and quality of the good. Various surveys exist getting at the first issue, but the second (regarding the quality of counterfeited goods) is one that has not yet been explicitly considered as part of the counterfeit literature (only in implicit terms of placing a dollar value on the replaced product). The fact that these may be uncertain parameters or parameters that change from product to product suggests that economic modelling, which applies average behavioural responses to limited and uncertain data, may be very useful for improving estimates in the future.

The actual quality of a pirated product is often known with more certainty by the consumer and it is usually similar or close to that of the legal product. Hence, the more important factor one needs to understand for obtaining a UUPC rate for a given product is the average consumer's willingness to pay for a pirated good (which might differ based on the baseline price of the legal good as well as the legal risks). This is fundamentally different from counterfeited products, where the quality of the counterfeit good is more likely to be inferior (and possibly unknown) to the consumer, and where the status that comes with original products (particularly luxury brands) is important.

There is a lack of clarity on methods for extrapolating to other markets or countries
In the literature on both counterfeiting and UUPC there is lack of consistency and often clarity regarding how to extrapolate for missing data, whether it is missing based on market segment or national data. Here a common treatment of how to deal with missing data across the literatures is probably feasible if it is based on the relative size of a given product market, level of aggregation across nations, or some other basic feature related to the aggregation of the product across markets/countries. This is not to say that we are recommending the same methodology be applied for each of these types of estimates (product-based, nation-based, region-based), but rather that the general treatment of the problem (missing data at the region level, missing data at the market level, missing at the product level) can be handled in a standardised fashion regardless of whether the product pertains to counterfeiting or UUPC. More serious considerations regarding how to aggregate findings for specific products across countries need to be contemplated in the light of

the nuances of the different product markets. While aggregate studies of UUPC and counterfeiting are indeed useful, broad assumptions of stable rates of UUPC and counterfeiting across product types are not reasonable or reliable.

The trade-off between accuracy and coverage can be initially resolved in favour of the former
In the case of counterfeiting and UUPC, it may be best to concentrate, at least initially, on specific products and/or markets, thus sacrificing some coverage. This is because the hidden nature of these markets requires that significant effort and learning is required when trying to measure these phenomena in a meaningful way across countries and product groups. Significant effort is necessary before trying to repeat the approach for other markets in every country (as some countries will certainly have worse data than others). In other words, in a hypothetical production function for a measure of counterfeiting and UUPC, higher weight is attached to the input accuracy (especially at this relatively early stage at which there hardly exist reliable benchmarks). This speaks in favour of an approach to the estimation of the size of counterfeiting and UUPC that identifies 'case studies' (i.e. specific markets and/or countries) that are more immediately tractable due to intrinsic characteristics or availability of data. A methodology could then be pilot-tested in these cases. One example of this approach can be found in the OECD 2008 study, whereby the data requirements for the construction of one of the indexes of counterfeiting and piracy (the ATRIC) led to restricting the scope of the analysis to four economies only in order to reach the necessary level of detail. Appendix C provides more detail on how we envisage this approach, and the steps we believe are necessary to undertake it.

The overlap in the literature between size and effects reveals a preference for "market-based" approaches
As discussed above, many studies focus on sales lost to legitimate IPR holders, which can be considered as a measure of the size of counterfeiting and UUPC, but also represents the first order of their effects. Clarifying the distinction between size and effects from a conceptual perspective is not a crucial question for future efforts in this field. However, from a practical standpoint our review suggests that lost sales, measured in terms of quantities or revenues, are a sensible outcome to consider when estimating the size of counterfeiting and UUPC using a 'market-based' approach, i.e. an approach that somehow tries to model demand and/or supply. Other approaches require data that are not available, harming the consistency of the estimation process. For example, relying on seizures without knowing the effective detection rate is problematic. The estimation of the second order of the effects of counterfeiting and UUPC (i.e. the effects on governments, such as tax revenues, and consumers, such as health and employment) is far less problematic in terms of data requirements (once a consistent estimate of the size of counterfeiting and piracy exists).

A first-cut quantitative analysis of the effects of counterfeiting and UUPC does not require complicated methodologies
Given the current state of knowledge, studying the impact of counterfeiting and UUPC does not require fancy econometric techniques. The linear regressions found in the literature so far are a good starting point provided that the right interpretation is attached to them. The biggest challenge remains in obtaining reliable measures of the right-hand-side variables, i.e. factors that correlate with the magnitudes of counterfeiting and UUPC. While these are being developed it would be valuable to compile a body of knowledge from currently available studies, which can guide future empirical analyses: a comprehensive taxonomy of the types of effects of counterfeiting and UUPC that many studies have qualitatively described, in order to assess for

which ones there exist the data (i.e. the measures of the outcomes, the left-hand-side variables) and possible methodologies (econometrics, simulation, etc.) to attempt a quantitative estimation. As highlighted earlier, a significant improvement over the existing literature also requires the inclusion of consumer surplus.

Any methodology should be 'triangulated' with other methodologies
Estimates from this new method should also be compared (or triangulated) with counterfeit and UUPC estimates generated from alternative methodologies, either from studies conducted by firms using various industry gold standards or by outside parties. We expect a role of such external validation with estimates from different methodologies at two different stages: 1) during the further development of the methodology, and in particular during the pilot testing; 2) periodically after implementation of the methodology, for example every three or five years.

In the short-term external validation efforts should occur when pilot-testing the methodology with a particular firm. It forms a basis for discussion of applicability, helps understand potential weaknesses of either method, helps identifying potential improvements and finally (if the results mimic one another) it helps in building trust.

In the longer term, triangulation efforts with other methodologies may occur when the method is implemented, and the model is run by OHIM every year with firm data. At this later stage, it would be useful, if not necessary, to carry out a periodical triangulation of the method with an alternative ('gold standard') method to identify potential developments that may not be captured by the new approach. These latter efforts are typically much more resource-intensive and would therefore only be employed (for example) once every five years, and only for a limited number of products.

4.6   **In sum**

As counterfeiting and UUPC remain high in the policy agenda, there are continued attempts to present a quantified picture of the situation. While counterfeiting and UUPC are a business concern across all the industries affected, counterfeiting of safety-critical goods (such as pharmaceuticals, car and aircraft parts, and pesticides) may be particularly worrisome if they can lead to serious health harm or loss of life. In this context, business, governments and society as a whole have a stake in quantifying the extent of this problem. This chapter makes an important contribution to these efforts. To our knowledge, no other paper has attempted critically and comprehensively to review existing measurements of the scale of IPR violations, although some have noted weaknesses in this body of research.

Based on this rigorous and extensive literature review, we conclude that there is no reliable or accepted method for estimating the size of counterfeiting and UUPC that is feasible for the purposes of producing annual measures for the all affected products/markets and in all countries. While different approaches (such as surveys of consumers and producers) can provide useful insights towards understanding specific aspects of these markets, it appears that no single approach can accomplish a comprehensive estimation of scope, size and impact with limited resources. The 2008 OECD report, which has been one of the more rigorous contributions to the field, itself argues that while it has been able to provide insights into the situation, 'the information base needs to be further strengthened'.

While it would be desirable to accomplish a comprehensive assessment of the scope, scale and impacts of counterfeiting and UUPC with a single study, we believe it is more valuable, given our current knowledge, to focus our attention on the single components of these complicated issues than on the bottom line. Therefore, we propose a 'building block' approach which recognises that innovation, growth, competitiveness, creativity, culture, public health and safety, employment, environment, tax revenues and crime are potential consequences of an underlying phenomenon which is very difficult to measure. Under this view, the assessment of these broader impacts cannot advance unless the measurement of infringement itself has reached a scientifically satisfactory stage: we cannot build a higher-level analytical layer unless its inputs are solid. Based on this principle, we believe that it is of foremost importance to focus on 'building the foundations' for a methodology that is consistent, robust, feasible and reliable in measuring the size of counterfeiting and UUPC. In a second phase it would then be sensible to work on improving the methodologies that are currently applied to studying the broader impacts of infringement. As discussed above, some of these methodologies would not require very sophisticated efforts to yield reasonable first-cut assessments of many of these effects; but such assessments can be useful only once a reliable measure of counterfeiting and UUPC is available that is comparable across countries and sectors and over time.

To advance the field in this direction, our proposed approach builds on economic theory to derive relationships that link observable quantities in the market to the presence of counterfeiting and UUPC. This represents a novelty with respect to the existing empirical efforts in this field because we take this one step further by proposing the aggregation of firm-specific information on unexplained variation in sales to generate a global estimate of counterfeiting. Other efforts to estimate empirically the total size of the market for IPR infringements make no or little use of the simple standard theories that model the decisions of the economic actors involved.

Estimates from this new method should also be compared (or triangulated) with counterfeit and UUPC estimates generated from alternative methodologies, either from studies conducted by firms using various industry gold standards or by outside parties. We expect a role of such external validation with estimates from different methodologies at two different stages: 1) during the further development of the methodology, and in particular during the pilot testing; 2) periodically after implementation of the methodology, for example every three or five years.

Our proposed approach is introduced and described in detail in the following chapter.

CHAPTER 5 **Testing a new approach for estimating the scale of IPR infringements**

In Chapter 4 we concluded that while most of the methods that currently exist to estimate trends in IPR violations are informative with respect to understanding infringements, each of these measures has drawbacks related to data selection, costs of collecting the data or lack of information on substitution rates. In our work we propose and test a new methodology for estimating the total size of the counterfeit market across all products. This theoretical framework, motivated by economic theory, has been applied to estimate impact of infringement on revenues in specific industries in a few instances (for example: Qian, 2011). We propose to use forecasts from firms combined with information in the literature on country-level measures related to counterfeiting to understand counterfeiting trends. Our key insight is that the relationship between these country-level measures and unexpected differences between firms' forecasts and sales provides us a method for estimating trends in IPR violations. In this chapter we highlight the findings of an empirical test of this new methodology using confidential data made available to us by an industry partner producing consumer goods targeted by counterfeiters.

Whilst the theoretical framework underpinning this methodology should in principle apply to UUPC as well, there are a number of sector specificities that make it difficult to pilot-test in this area. As the empirical data used in this chapter refer only to products targeted by counterfeiters, we focus on this type of IPR infringement here. Chapter 7 discusses the implications of these sector specificities for extending the method to unauthorised use of (online) protected content.

5.1 Theoretical underpinnings of RAND's model

The methodology we develop to estimate the size of the market for counterfeit goods is a supply-side approach making use of economic theories of monopolistic competition and differentiable goods. The idea is to estimate from legitimate producers the post hoc amount of 'unmet demand' that they experience and use that as a proxy for total counterfeits sold. We presume that legitimate producers who are targeted by counterfeiters are able to calculate this amount as the difference (or 'residual') between their forecasted demand for their products and actual units sold.[11] We are able to make use of information from legitimate producers to determine back out information on unmet demand under the premise that authentic products are targeted for two main reasons:

---

[11] One criticism of this approach is that the quality of these estimates from firms may be inadequate, a point we will speak to later in this chapter.

1) They have a known, branded, desirable (and probably differentiable) product that:

   a) because of these attributes, can command supra-economic profits for legitimate producers (that is, they receive a per unit price that is above the actual cost of production and hence each unit sold generates a positive amount of economic profit); or

   b) because of costly inputs to production (unique technology, highly skilled labour, payment of royalties, government-limited production, licensing regulations), can be illegally produced at lower (or no) cost (and hence produce a profit for counterfeiters).

2) The product is sold in a market in which the distribution of the good can be penetrated by:

   a) counterfeiters using the legitimate distribution system for the authentic product and trying to compete via normal retail channels (i.e. the 'primary market') with buyers (presumably) unaware of authenticity; or

   b) individuals creating an alternative 'secondary' market place either in a physical location (street market, selected 'friendly' retailers) or on the web, where consumers can purchase with some knowledge that the authenticity of the product is in question.

We note that the vulnerability of a firm or product depends on firm- and/or product-specific factors (such as the quality of the good and the ease with which a counterfeiter can replicate that quality) as well as a variety of general market factors (such as the legal environment for IPRs; the level of enforcement by legal entities and targeted firms; and social norms surrounding that product are more or less accepting of non-authentic goods).[12] While these are clearly important, and our model will attempt to consider them in an alternative way, the construction of the model is based on market structure (which we refer to as 'types'), not these factors.

By and large, there are two types of markets: competitive and non-competitive. Competitive markets are characterised by the fact that they include a large number of buyers and sellers; there is perfect information on the cost of production and products sold; and there are no barriers to entry or exit. In perfectly competitive markets, no single buyer or seller can affect the price. The price of a good or service is set by the market and firms simply take the price as a given and decide on how many units to produce given that price. The implication, therefore, is that if a firm in a perfectly competitive market tried to increase the price of its product or service, all their consumers would purchase from other firms and that firm's sales would fall to zero.

In monopolistic competition, the industry is still characterised by a large number of firms possessing perfect information about each other's products and production processes and there remains free entry and exit from the market. However, in these competitive markets, firms are able to differentiate their products from other firms, so the firms provide goods that are similar in function but have slight differences that are either real (physically different or with different support services attached) or perceived. The ability to differentiate the product of one firm from other similar products sold in the market allows individual firms to charge higher prices than in perfectly competitive markets, but not substantially higher because of the availability of substitute

---

[12] See for example Vida (2007).

goods.[13] The ability to influence price for their own good by choosing specific output levels is what sets competitive monopolistic industry apart from perfectly competitive markets.

In non-competitive markets (oligopoly or monopoly markets), there are relatively few firms providing the total output to the entire market and often significant barriers to entry exist (due to patents, ownership of unique resources, specialised technology or advertising budgets). Since relatively few firms provide the total output to the market (only one firm in the case of a monopoly), these firms face the entire market demand curve for a good and hence can influence the overall price of the good sold by changing their output (and hence total supply). If a firm with market power in a non-competitive market increases the price of the product, the quantity of goods sold by that firm remains positive but is smaller than that which could be sold at a lower price (due to the law of downward sloping demand, and hence diminishing marginal revenue). This of course depends on the degree to which consumers are still willing to pay for the product and how sensitive they are to changes in price (i.e. price elasticity of demand). [14]

Counterfeiters exist in both monopolistically competitive and non-competitive markets. However, their interest in competitive markets only exists in those markets where products are successfully differentiated, as those are the products that can receive monopoly rents for at least a short period of time. The more successful a firm is at differentiating its product, the greater the incentive for counterfeiters to enter that market.[15]

The fact that counterfeiters are more attracted to markets where firms are able to extract some monopoly rents (either due to product differentiation or because it is a true oligopoly) is precisely the feature of these markets that we attempt to exploit in our empirical strategy to estimate the size of the market. The logic is most simply presented using the example of oligopolistic markets where it is presumed that a firm or small number of firms have significant market share.

In an oligopolistic market we can presume that the change in production observed from the main leader would generate a measureable change in total supply to the market, as the market leader by definition influences total supply and the supply response of the other firms (Mansfield, 1993). Because of this, it acts as if it faces the market demand curve, or, more accurately, some known proportion of it. Conditional upon assumptions of its competitors' behaviours, it will set quantity (and hence determine price in the market) at a level that maximises its own profit given the known market demand. To the extent that the market leader does a reasonably good job

---

[13] The implication is that the firm faces its own "market" demand curve, but it is fairly elastic. The more available are good substitutes, the more elastic the demand curve the competitive monopolist faces. Therefore, they can only vary price up to a point, beyond which all consumers switch out of that market.

[14] 'The concept of price elasticity of demand is used in economics to describe the sensitivity of consumption to changes in the monetary price of a product (i.e. the percentage change in consumption resulting from a 1-percent increase in price). For example, a price elasticity of alcohol demand of -0.5 means that a 1 percent increase in price would reduce alcohol consumption by 0.5 percent" (Hunt, Rabinovich and Baumberg, 2011, p.7).

[15] As described in Qian (2011), counterfeit sale of goods is an endogenous process where the more the authentic producer sells, the greater the likelihood of counterfeiters copying the brand. As such, there is a positive correlation between the number of authentic products sold and the number of counterfeits on the market (i.e. the more successful is a brand, the more counterfeiting there is). It is not necessarily the case that counterfeiting causes increased demand for authentic products, although it can occur in that counterfeit and pirated products can inform consumers about the authentic good and increase demand for the authentic good (Shapiro and Varian, 1999; Conner and Rumelt, 1991).

predicting the other firms' behaviours (which are completely observable after the fact), it should do a very good job predicting what it can sell itself.

If a leading firm in an oligopoly market is unable to sell the predicted quantity it projects, it is typically due to an unexpected shock that is observable after the fact, such as a shock in the supply chain, poorly received advertising campaign or even a financial crisis or earthquake that impacts the overall economy. Once these factors are taken into account, the revised predicted quantities look a lot more like the volumes actually sold. However, sometimes the revised projection (accounting for post hoc shocks to the product or market) still cannot account for the difference between revised forecasted sale and actual sales. This 'unpredicted forecasting error', which cannot be explained by normal observable product, firm or market factors, represents the error in prediction by unobservable factors. One of the unobservable factors is counterfeiting. Therefore, knowledge regarding the magnitude of the unpredicted forecasted error (i.e. the number of units that cannot be explained post hoc by observable factors) can be used to quantify the impact of counterfeiting on the sales of the authentic product, once we know how much of that unpredicted forecasting error is due to counterfeiting.

A very similar approach can be used in the special case of monopolistic competition as well because firms are able to differentiate their product sufficiently enough to generate their own demand. To the extent that firms make use of the information they have regarding their market demand to generate forecasts of unit sales and they can observe competitors' behaviour in terms of sales post hoc, then the exact same approach can be applied.

The fact that there are substitute goods also available, of varying quality, means that a firm's expectation of sales depends on the behaviour of firms that provide close substitutes (their marketing strategies, new releases of improved products, etc.). When actual sales do not meet the expected sales for an original firm selling a differentiated product, the potential explanation is more than just unexpected shocks in market demand as it may be that a known competitor was successful in pulling away consumers from the original firm. Thus, it will be important for developing estimates of the unexpected forecast error to understand observable shocks in these competitive markets to understand whether there were changes in the distribution of sales across substitute goods before inferring the amount of forecasting error that can be truly attributable to 'unexplained' factors like counterfeiting.

5.2   Empirical strategy

Relying on the assumption that counterfeiters are only interested in entering markets for which there are some sort of rents, we can adopt the following two-stage empirical strategy to identify the amount of unexplained forecasting error that is caused by counterfeits and aggregate that up across markets to get the total impact. We consider for now the case of non-competitive markets (specifically, an oligopoly) where a firm is able to set its own price.[16] Based on this price, it determines the quantity supplied to the market, meeting either the entire market demand for its product or its anticipated share based on that price. Thus, variation in price/quantity, in the absence of counterfeits, would only occur because of one of the following occurring:

---

[16] The model can be similarly applied to a monopolist or discriminating monopolist producing a differentiable good.

- firm factors shifting the cost of production unexpectedly and hence the profit-maximising output

- demand-side factors shifting the market demand (e.g. changes in disposable income, new product substitutes or complements entering or leaving the market)

- market factors altering the structure of supply or demand (e.g. a natural disaster affecting the supply-chain; changes to government regulations).

If these factors are held constant and the quantity of product a firm is able to sell (at its selected price) is below the expected amount, then this is an indication that a counterfeiting product has entered the market and taken away some of the market demand. In other words, the amount of counterfeit product can be 'backed' out by considering how much the oligopoly or monopolistic competitive firms expected to sell and how much they actually sold.

Thus, we can estimate the model empirically in two stages. In the first stage the proportion of the forecast error that can be explained by these ex-post unanticipated market- or industry-specific factors is estimated so as to obtain an estimate of the unexplained forecasting error. More formally, we expect that each firm can estimate some form of the following equation:

$$q^*_{ijkt} = \beta_1 x_{ijt} + \beta_2 y_{ijt} + \beta_3 z_{jkt} + \varepsilon_{ijkt} ,$$

(1)

where q^*_{ijkt} is the total forecast error at time t for firm i in a market (which we define as by product j and country k). x is a matrix of product-specific variables that were not anticipated so are not included in the original forecast; y is a matrix of firm-specific variables that reflect unexpected changes in relevant firm factors, and z is a matrix of market- and demand-level variables that might contribute to a forecast error (and is therefore product-country specific, jk, not firm-specific).

The regression specified in equation (1), therefore, includes as regressors only those things that the firm did not accurately know when developing its original forecast for the period, but learned post hoc. The coefficients $\beta_1, \beta_2, \beta_3$ tell us the relative importance of these factors, firm and market shocks, respectively, in predicting the error. ε_{ijkt} represents unobservable factors that also influence the forecasting error, which might include random noise but would also include unobserved fluctuations in the amount of counterfeiting/pirating going on in the particular product market.

By estimating equation (1) empirically, each firm can generate an estimate of the predicted residuals ($\hat{\varepsilon}^*_{ijkt}$) which quantifies the unexplained variation in the forecasting errors (i.e. those not caused by unpredictable changes in x, y or z that were observed at the end of the forecasted period). It is this unexplained variation in the forecasting error that is then used in our second-stage model, which aims to understand what fraction of the unexplained variation in authentic product sales is accounted for by counterfeiting.

Of course, the key to being able to use this market logic to estimate the size of the counterfeit market depends on having good information on which to base expected sales, holding actual market conditions constant (or, more accurately econometrically accounting for them). The current assumption that firms will have this information is based on an understanding that they

need to be able to have an appropriate amount of inventory available and plan resources (e.g. labour costs, interest payments, etc.). For this reason, firms make forecasts about how much quantity they expect to sell. In practice, this may be done through a variety of approaches including simple approaches of considering past trends in sales and sales of similar products, to more sophisticated econometric techniques accounting for a variety of factors. [17]

Regardless of approach, this first stage estimation involves highly sensitive firm-specific information (about their forecasts, tracking of competitor sales, etc) that most firms would be reluctant to share with a centralized analyst on a regular basis. The unexplained forecasting error, which is all that is needed by a centralized analyst to estimate the second stage model using multiple firms data, is presumed to be less sensitive. This is because the number represents a compilation of a variety of different types of information, and as such masks a lot of the (potentially) commercially sensitive information. The unexplained forecasting error should be fairly low vis-à-vis actual sales, once the other factors are taken into account and hard to interpret by itself. Thus, it is presumed that if this approach were to be broadly adopted by the European Commission or an agency on its behalf, firms might implement the first stage using their own data and merely provide to a centralized agency the less sensitive information regarding how much of their forecasting error could not be explained by systematic market or firm things they track. Given this assumption, the basis for the two stage estimation is to mimic the process that could occur if implemented in a real world setting.

The second stage of the empirical process is to assess what proportion of the remaining forecasting error can be attributable to counterfeiting. In a second regression, we estimate the relationship between the 'unexplained forecast error', $\hat{\varepsilon}^*_{ijkt}$, for firm i operating in product market jk at time t and factors of counterfeiting in those same markets. Formally, the second stage regression is of the following form:

$$\hat{\varepsilon}^*_{ijkt} = \beta c_{jkt} + u_{ijkt} , \tag{2}$$

where $\hat{\varepsilon}^*_{ijkt}$ is the amount estimated previously for the unpredicted forecast error of a product by firm i at time t and c is a matrix of variables related to counterfeiting that contribute to the unpredicted error in forecasting (discussed in greater length below). Firms may over- or under-sell the amount forecasted because of reasons completely unrelated to counterfeiting and for market factors not yet taken into consideration by the model;[18] this is captured by the new error term u_{ijkt} .

---

[17] We recognise that not all firms may be able to estimate a formal model to generate their unexplained forecasting error measure, which will lead to an overstatement of the unexplained forecasting error in the second equation. Assuming that variation in counterfeiting is not systematically linked to unexpected changes in firm-, factor-, or market-characteristics, the use of a grosser measure of unpredicted forecasting error would just add noise to our second stage of the model.

[18] In practice, this can mean demand out-stripping supply, and firms needing to make additional purchase orders to the normal purchase schedule.

Upon estimation of equation (2), one can generate a prediction of the conditional mean, $\hat{\beta}c_{jkt}$, which is a direct estimate of the amount of unexplained forecasting error that can be predicted by counterfeiting supply and demand factors:

$$\hat{\beta}c_{jkt} = \text{Amount of counterfeiting} \qquad\qquad\qquad (3)$$

Appendix B provides a non-technical description of the methodology and a series of common questions and answers. This description was sent out to potential participants in the pilot phase.

5.3   Potential threats to our identification strategy

While the empirical model just described stems from economic theory, there are at least three obvious threats to the identification strategy we propose here that need to be recognised and considered.

5.3.1   Quality of forecast data from firms

First, the strategy presumes that firms actively engage in forecasting and can reliably identify factors that go into their forecasts, understand how random shocks might influence their actuals and have the data and capability to go back and assess how they would have done had it not been for those realised shocks. This is possibly a very tall order, despite the fact that most of the firms targeted by counterfeiters are companies that have shareholders or private investors they must report back to. Even if firms can do a reasonably good job of explaining their own firm factors that lead to forecasting errors, they may not do a very good job tracking competitors' behaviours, which the model suggests are an important factor to consider. To the extent that firms do a reasonably poor job of understanding why forecasts are not met, or to the extent that firms lack data enabling the quantification of these issues, the first-stage model might be poorly specified.

The implication of the poorly specified first stage depends on the extent to which counterfeiting measures included in the second stage of the model are correlated with firm, industry and market factors relevant for understanding forecasting errors outside of counterfeiting. If the measures are uncorrelated with these shock factors, then the misspecification in the first stage should have no impact on the quality of the estimate obtained from the second stage. The counterfeiting variables will explain the same amount of variation in the error term in the second stage regardless of whether the first stage can be estimated if these variables are independent from first stage shock factors. If, however, there is a positive or negative correlation between the counterfeiting measures and some of the factors used to develop forecasts, then the results from the second-stage model will be biased when the first stage is misspecified. The direction of the bias is not possible to sign a priori and will depend on the direction of the correlation between the forecasting variables and the counterfeit measures.

It is possible within the context of our current pilot to assess whether a correlation exists by estimating a variant of our two-stage model, which we refer to as a 'simplified' model. This basically ignores the first stage and goes directly to estimating a second-stage model using our counterfeit measures as controls and the simple difference in forecasts and actuals as the dependent variable. Doing so allows us to assess, at least for this single firm, what the implication would be of ignoring firm-, industry-, and market-specific factors that are important for developing forecasts but unavailable to an outside analyst for whatever reason.

5.3.2   **Forecasts which incorporate an element of counterfeiting**

A second but related threat to identification exists if a firm is not able to identify the extent to which its forecast already includes an estimate of counterfeiting. Such a situation might occur if the firm simply uses the last year's actual units sold to forecast what it expects to sell in the next year. The threat occurs in this instance because firms are unable to identify to what extent counterfeiting is already incorporated into their forecast for the next year (as there is no model on which a forecast is based and which can quantify counterfeiting). To the extent that lost sales due to counterfeiting are already incorporated into projections for the next year, then the second-stage model will only be able to capture deviations from the baseline value presumed in the forecast. It will, therefore, systematically underestimate the amount of counterfeiting occurring in the primary market for this firm's product. While this will obviously generate a systematically undervalued aggregate estimate of the level of counterfeiting, it will not necessarily systematically bias trends in counterfeiting generated from the model, as the model will still pick up deviations from the mean value, which by definition will influence the trend over time. This is an important distinction, as the model is anticipated to underestimate counterfeiting overall anyway given its inability to measure the amount of counterfeit goods sold in the secondary market (and not competing with the authentic firm). But even though the model will underestimate counterfeiting in total in a given year, it could still generate reliable information on how counterfeiting is changing from year to year.

5.3.3   **Omission of an important variable in the second stage**

A third potential threat to our identification strategy is the unintended omission of an important variable from the second-stage regression (or general misspecification of the second-stage model). Again, the potential impact of an important omitted variable critically depends on the degree to which it is correlated with an included variable.

The usual concern is that an omitted variable that is correlated with an included variable will generate a biased estimate of the marginal effect of the included variable on the outcome of interest (in this case counterfeited units). This is because the included variable will indirectly capture the influence of the omitted variable, to the extent that they are correlated. In the current instance, we are less concerned with accurately measuring the marginal effect of specific counterfeiting measures in the second stage, but more concerned about generating a reliable estimate of counterfeiting (or trends in counterfeiting over time). Thus, the omission of an important variable, if it is correlated with an included variable, might still be partially represented in the predicted value of counterfeited units because of its relationship with the included variable. The bias of its exclusion will be less than if it is completely independent of all the included variables.

One of our strategies to deal with the problem of omitted variables in the second stage is to include country fixed effects in this second-stage model. Doing so allows us to capture time-invariant unobservable factors, such as norms regarding willingness to buy counterfeited goods, for which we do not have direct measures. As the time period we are analysing is relatively short (only six years), it is unlikely that these sorts of unobserved country-level factors will be changing much over time, so the inclusion of the country fixed effects allows us to capture their influence on predicting the number of counterfeit units sold indirectly. However, to the extent that the omitted counterfeiting variables are not measured at the country level or are not time-invariant, the model will generate predictions that are biased upward or downward (depending on the relationship between the key omitted variable and the total level of counterfeiting).

While it is clear that we cannot perfectly fix all of these issues of identification, it is our belief that the inclusion of firm- or product-specific fixed effects in the first stage and country-specific fixed effects in the second stage will mitigate the influence of many of them. Of course, it is only through comprehensive testing of the base model using alternative model specifications that one can be certain that our approach handles these issues reasonably well. Such a thorough assessment will be possible only if data from a large number of firms is available on which to test the model's assumptions.

5.4 Data

It was our goal to estimate the model described above using data from multiple industry partners, so that we could test the viability of our proposed methodology across various products and industry sectors. Unfortunately, we were able to obtain data from only a single industry partner. Thus, the empirical assessment of our proposed methodology is limited and should be viewed more as suggestive rather than definitive evidence of the viability of this approach. While we believe that this approach can be modestly modified so as to include data from multiple industries and products, we leave such evaluation for future work.

A single industry partner, whose name shall be kept confidential, provided proprietary historical data for the period 2006–2011. The technology firm is an internationally renowned leader in the sale of a specific technology offered through a variety of products patented by the firm in numerous countries all over the globe. Because of the success of its products and its patented technology, it has been the target of counterfeiters for over a decade. The firm's own interest in understanding the impact of counterfeiters on its sales and profitability has led it to invest substantial money into measuring the size of counterfeiting. In fact, the firm currently generates estimates of the impact of counterfeiting on its bottom line using mystery shopping methods for a sample of cities and then extrapolating those findings to other cities and markets. Its interest in participating in this project with us was to assess how our estimates compared with its own estimation strategy, considered the 'gold-standard' by many, given that it is not possible for the firm to carry out mystery shopping activities in all markets or on an annual basis.

The firm provided us with information on over 50 related products sold in more than 30 countries across six global regions. Data on firm-specific variables used to extrapolate counterfeit estimates in years when mystery shopping is not possible were provided, in addition to information on forecasts and actual units sold during a fiscal year. It was not made known to us, however, in which years the mystery shopping occurred and which years were model extrapolations (or which markets were pure extrapolations). Key information used by this firm in the construction of its forecasting estimate included: information on the install base of a particular group of technologies making use of its specific product; the characteristics of the product itself (some characteristics were more desirable than others from a market perspective); and sales of key competitors of similar products in specific markets. Forecasts were constructed quarterly by country, so we relied on data obtained from the first quarter of the calendar year.

Not all of the products for which we were provided information on forecasts and actual units sold contained the firm-specific product and market information for the years of interest. Thus, while the final data set used for testing our model contained information on 45 related products sold in 16 countries over the period 2006–2011, the actual data set included only 3,166 observations

due to missing values for some products/years/countries. To these firm-specific data, we added a measure of the general market demand in each of the markets in which the firm was operating, using a measure of the rate of growth of the economy. Information on the growth of GDP for each country included in our analysis was obtained from the IMF.

For our second-stage regression, we merged into the existing firm data some measures of the susceptibility of these products to IPR infringement, which are obtained from World Bank surveys. The World Bank collects national-level data on a series of topics systematically from countries across the globe, including economic prosperity, trust in the legal system, and so on. For our study, we drew on three indices constructed by the World Bank to capture elements of the political and regulatory environment that research suggests are associated with greater/less counterfeiting, namely indicators of:

- the rule of law;

- control of corruption; and

- government effectiveness.

Ex-ante we expect that the rule of law and control of corruption variables are more likely to related to counterfeiting than the government effectiveness variable, but in this pilot exercise we experiment with all three variables.

The World Bank defines its rule of law variable as a measure of survey respondents' perceptions about how well the rules of society are abided by. Variables used in the construction of this index come from surveys of individuals who report on their beliefs regarding the effectiveness of the police, confidence in the policy system, whether intellectual property protection is weak, speediness of the judicial system, enforceability of contracts and trust in the functioning of the criminal justice system. The control of corruption variable, as defined by the World Bank, captures peoples' perceptions of the extent to which public power is used for private gain. Questions used in the creation of this index include the frequency with which firms are required to make payments in a variety of settings (favourable judicial decisions, public utilities, etc.), the frequency of corruption amongst public institutions such as the state legislatures and customs, and the existence of country anti-corruption policies. The government effectiveness measure attempts to assess how respondents feel about the quality of their public services, civil services, policy formation, and independence from political pressure.[19]

We also consider as additional variables in our second-stage analysis two indicators of the complexity of customs procedures, namely the number of documents required to import goods (documents) and stringency of a country's customs procedures (custom's burden). In the analyses that follow we only use the measure of custom's burden, as we found the documents measure to be highly correlated and provide no new information. Furthermore, we include two country-level measures that may be tied to the level of demand for counterfeits: international tourism and broadband access. Both measures are hypothesised to be positively related to the level of counterfeits (and hence unexplained differences in units sold). The source of all of these data is the World Bank's Development Indicators.[20] For the data used from this one firm, neither of

---

[19] The full set of questions related to the construction of each of these indices can be found on the Collaborative Governance page of the World Bank Institute website: http://wbi.worldbank.org/wbi/topic/governance.

[20] As of 20 March 2012: http://databank.worldbank.org

these variables added independent variation to the second-stage model; they too were highly correlated with other variables, which is likely a function of the limited variation in product types or the focus on only one firm. Future efforts drawing on data from other firms or product groups should still explore the utility of these variables.

Descriptive statistics on some of the key variables of interest for our analysis are reported in Table 5-1. Of course, specific details relating to the firm's products that might lead to identification of the firm have been withheld. In the first-stage regression, we estimate the difference in forecasts and actuals (the third variable listed in Table 5–1) using predictors specified in the second section of Table 5–1 (labelled 'Independent variables – 1st stage'). In the second stage of the model, we use the predicted unexplained forecasting error (which is the predicted error from the first stage) and use variables specified in the bottom portion of Table 5–1.[21]

Table 5-1:   Descriptive statistics of firm data

Dependent variables		
	Mean	Std dev
Firm forecasts (units sold)	708,957	1,941,571
Firm actuals (units sold)	695,984	1,977,649
Diff: forecasts – actual	13,512	222,772
Predicted unexplained forecast error (from 1st-stage mode)	0	197

---

[21] It should be noted that the empirical model specified in equations (2) and (3) of the previous section suggest that the error term will capture variation in different product markets ('jk'). If we were estimating a model with multiple firm data from different product groups, that would be the case here. However, because all of our data come from a single firm and product group, we are unable to differentiate a product market ('jk') from the general market ('k').

Independent variables – 1st stage		
	Mean	Std dev
GDP annual growth	2.9	3.86
Existing base technology previously sold	965,990	1,888,199
Competitor sales product type A	191,547	423,684
Competitor sales product type B	96,038	172,832
Competitor sales product type C	24,680	36,853

Independent variables – 2nd stage		
	Mean	Std dev
Rule of law	0.94	0.8
Control of corruption	0.88	0.9
Government effectiveness	1.01	0.63
Custom's burden	4.48	0.52
International tourism	7.58	3.90
Number of documents (dropped)	5.39	2.11
Broadband access (dropped)	20.18	9.76
N	3,166	

Notes: Product data are from countries within North America, Europe, Central, South and East Asia, and the South Pacific. Spans the 2006–2011 time periods. Data captures 45 unique products across the 2006–2011 time period. Missing values are imputed for policy variables so as to retain all countries in which we have firm data.

5.5   Results

The two-stage model described above was estimated using data from a single firm selling 45 different (but related) products in 16 countries. Coefficient estimates from both the first- and second-stage models are provided in Table 5–2.

First-stage regression
Key insights from the first-stage regression are the following:

- Actualisations of several firm- and market-specific factors used to generate forecasts by the firm are indeed useful for explaining 23 percent of the unexpected difference in forecasted units sold and actual units sold. Of course, the firm may have additional information at its disposal that could be included in the regression and increase the predictive power, but the bottom line is that about a quarter of the difference in forecasted units and actual units sold can be explained with information collected retrospectively.

- The proportion of products sold with a particular characteristic that is commonly targeted by counterfeiters has larger forecasting errors, which is consistent with findings from Qian (2011) that the volume of authentic goods sold is positively correlated with the volume of counterfeits sold.

- The firm does a better job forecasting in markets where there is more base technology making use of their products, which we interpret as indicating that the firm is better at forecasting in markets that are more established and not experiencing large growth.

- Product fixed effects, which are also included in the first stage but suppressed from the table, as a group are important controls for explaining differences in forecast and actual units sold even after market factors are taken into consideration. Given our ability to control for competitors' behaviours and product characteristics, we interpret this as indicating important variation across products in the firm's ability to reliably forecast sales, which may have something to do with the product (it is new or growing quickly) or something to do with the markets in which the product is sold.

Second-stage regression

As the firm- and market-specific variables explained only a quarter of the variance in the unexpected forecasting error, there remains a sizeable portion of this difference that is unexplained. The question is, how much can our counterfeiting measures explain? The results shown at the bottom of the right side of Table 5-2 suggest that they explain some of the variation, but not that much. Only 5 percent of the remaining unexplained difference in forecasting errors could be explained by our counterfeit measures and controls in our second-stage regression. The fact that the counterfeiting measures explain a relatively small share of the remaining difference does not necessarily mean, however, that the model will do a poor job of estimating the number of counterfeit units. It does suggest that the estimates might be influenced by bias if important controls associated with counterfeiting have not been captured in our data and are hence omitted from the second stage.

Table 5-2:   Model results from two-stage estimation

Results from the First and Second Stage Regressions					
1st Stage Dependent Variable			**2nd Stage Dependent Variable**		
	Forecast Error			Residuals From	
	(in 1,000 units)			First Stage Regression	
GDP Growth	2.87		Rule of Law	235.80	
	(2.069)			(151.11)	
% Sold with Product Characteristic X	140.009	***	Control of Corruption	-306.17	***
	48.093			(101.670)	
Install of Base Technology	-2.89E-05	*	Government Effectiveness	-9.122	
	(0.000)			(107.48)	
Competitor Sales Product Type A	7.72E-05	**	Customs Burden	58.578	***
	(0.00)			(18.60)	
Competitor Sales Product Type B	2.52E-03	***	Intern'l Tourism	16.457	***
	(0.000)			4.51	
Competitor Sales Product Type C	5.25E-04	*			
	(0.000)				
Year	24.903	**	Year FE	Yes	
	(10.121)				
Country FE	Yes		Country FE	Yes	
Product FE	Yes		Product FE	Yes	
R-Squared	0.23		R-Squared	0.05	
Unique Products	45		Unique Products	45	
Cluster	Product		Cluster	Country	
N	3166		N	3166	

Notes: Forecast Error, the dependent varibable in the first stage is defined as the (Forecast-Actuals). Product data is from countries within the North America, Europe, South Pacific, Central Asia, South Asia, and East Asia regions; spans the 2006-2011 time periods. Standard errors are clustered at the level identified at the bottom of the table. Statistical significance is indicated as follows: *** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

Before discussing the specific findings with respect to the effect of counterfeit measures in the second-stage regression shown in Table 5–2, the reader is reminded that the second-stage regression includes country-level fixed effects. These time-invariant country variables are included so as to capture unobserved time-invariant differences across the countries in factors that we cannot adequately measure, such as consumers' willingness to buy counterfeit goods. To the extent that social norms and other important but unobservable factors are constant within a country, then country fixed effects are a great way of adjusting the model for these factors. However, because we have a limited number of countries and years in our data, it is also possible that by including these time-invariant, fixed country effects, we are creating some collinearity problems or bias with the counterfeit measures. In particular, we include a variable that does not vary much over the relatively short time period we are examining (just six years from 2006 to 2011, and fewer than that in some countries).

In looking at the descriptive statistics shown in Table 5-2, it immediately becomes apparent that the customs burden measure may be reflecting such collinearity bias, as it has a particularly low standard deviation within the sample period for the countries included. So, the positive sign observed on this variable in Table 5-2 may be reflecting this bias. Similarly, the positive and marginally significant coefficient on the rule of law measure appears, from additional sensitivity analyses included in Pacula *et al.* (2012), to be collinear with the corruption variable. Pacula *et al.* (2012) show that unexplained forecasting errors are lower in countries with a stronger rule of law, stronger control of corruption and higher levels of government effectiveness when each of these variables is entered individually. However, when all three of the variables are entered in the model simultaneously –which increases predictive power– the coefficients on the government effectiveness and rule of law measures change. The collinearity may be due to a true strong positive association between these variables, or it may simply reflect the selection of countries that are included in our analysis (based on where the firm's products are sold). With data from more firms and more products over time, this type of collinearity problem should be reduced.

There are a couple of other findings that may appear a bit surprising at first glance. In particular, we see that countries with higher levels of customs burden and more international tourism experience greater unexplained differences in forecasting errors. The positive finding for customs burden is consistent with the idea that countries experiencing a greater counterfeiting problem may be more likely to adopt burdensome custom procedures. Similarly countries with higher levels of international tourism are also positively associated with counterfeiting. Given the inclusion of country fixed effects, it may be the case that holding these other factors constant, a higher level of international tourism might support local counterfeit markets for two reasons: (1) tourists are easier targets for counterfeiters to deceive, as they typically purchase a good from the national market once only; and (2) local authorities may be more lenient toward counterfeit markets in tourist locations as a sort of marketing lever. Additional data, which would allow us to test the robustness of these findings, will be useful for determining whether these findings are robust across other firms and product groups.

Explaining the difference between the RAND model and firm estimates

While some of the findings in Table 5-2 are unexpected in terms of the estimated direction and significance of specific counterfeit measures, the real question is how well findings from the second-stage model predict levels or changes in the number of counterfeit products available on national markets. In Table 5-3 we provide summary measures of the average amount of counterfeiting estimated per product (in total across countries and years) generated by both the RAND model and the firm's estimates.[22] The numbers reported under the "overall mean" in Panel A represent the average amount of counterfeited units sold across products in which the firm provided an estimate of counterfeiting.[23] In looking at the overall mean level of counterfeiting per product, we can see that overall, the RAND model under-estimates the total number of counterfeits, as compared to the firm's estimate. If we look at the estimate of average

---

[22] The firm's estimates are based on a combination of mystery shopping in a number of specific markets, and a model to extrapolate these estimates to markets where no mystery shopping occurs. The firm did not reveal to us which estimates are based on mystery shopping and which are based on extrapolation.

[23] The estimates for the RAND sample are restricted to the same product, country and years for which the firm provides data

counterfeiting per product by year in Panel A, it becomes evident that the RAND model underestimates the average level of counterfeiting in four of the years (2006-2010) and then overestimates it in 2011. Importantly, the trend for the firm estimate is decreasing across all six years, while the RAND estimate decreases from 2006-2007, rises in 2008, continues falling in 2009-2010, and then rises again in 2011.

At first glance, the inability of the RAND estimate to track the firm trend in some years suggests the model is weak. However, further details about the data reveals why this may be. Both the RAND estimate and the firm estimate have considerable variability or a wide range (indicated by the rather large standard deviations around the mean). This wide range occurs for two predominate reasons:

1. Particularly for Panel A, the estimate is rather artificial, since it is derived from new products being tracked by the firm in different years or in new countries (with some products getting dropped over time). In other words, the model had to try to track the trend in counterfeiting when products would fall out of the sample and new ones would enter or enter different country markets. The inconsistency in products and countries over time can affect the average value from year to year in ways that are not reflective of true variation in trends of counterfeiting for a given product.
2. The products are grouped together to provide an overall estimate, yet, there is variability across the products (in the amount of counterfeiting for each) and for an individual product (in the amount of counterfeiting across countries). This sort of variability is expected and is indeed what the RAND model is trying to mimick.

To adjust for the variability mentioned in point 1 above, we restrict the sample so as to only include those 5 products that are systematically measured in the same 5 countries over time (see Panel B). While these 5 products still are not tracked perfectly in every country in every year, there is greater stability in the reporting over time. This provides us with some confidence that the model is indeed capturing product-specific factors and is a reminder going forward in data collection efforts that this consistency is needed.

Panel B of Table 5-3 shows estimates from the RAND model for these five products as compared to the firm's estimates. Again, one sees that on average, the overall amount of counterfeiting per product predicted by the RAND model is lower than that of the firm, and in some years it is lower by a lot (more than 90%). However, the data also show a general trend that is more consistent with that of the firm data, at least from 2006-2009, when both series are declining. Yet, in 2010, the firm data continue to decline while that of the RAND model starts to rise, suggesting that the model does not match to firm estimates as well in the later part of the period.

To adjust for the variability mentioned in point 2 above (i.e. the variability across products in levels might be influencing the averages reported and hence their trends), we assess the percent of counterfeiting of each product as a proportion of actual units sold (see Panel C).

We see in Panel C of Table 5-3 a couple of important differences. First, while the level of counterfeiting predicted by the RAND model underestimates counterfeiting compared to the firm estimates, the proportion of counterfeiting by each product is far more similar. In other words, there appear to be some products for which a substantial amount of counterfeiting occurs, but for which there is also a large amount sold. These large numbers clearly influence the overall level of counterfeiting the firm data projects in Panels A and B, suggesting that the RAND model

will systematically underestimate counterfeiting compared to the firm estimates. The evidence in Panel C suggests that the underestimation by the RAND model compared to the firm estimates might not be as large as originally thought, when considering that unit level data may be heavily influenced by a few particular products.

A second important difference from the RAND and firm's estimates made evident in Panel C is the difference in trends for counterfeiting suggested by the firm estimates. In Panel C, the firm estimate shows a rise in the amount of counterfeiting as a percent of actual units sold in 2008, but a decline when using level values (seen in both Panel A and Panel B). In Panel C, the decline resumes in 2009 and 2010 according to the firm data, but then increases in 2011. Importantly, the RAND estimates do not appear to capture the rise in counterfeiting suggested by the 2008 firm data (in Panel C), and instead show a continued decline and basic levelling off between 2009-2010, with a rise in 2011. In the third row of Panel C, we evaluate the full RAND predictive model, where estimates are predicted even for countries and products in which the firm did not provide counterfeiting data (although sales and projection data are available). We refer to this as the "RAND Full Sample (FS)." Here we see in this third row of Panel C, that the model predicts a trend for global counterfeiting of these products that looks more similar to that of the RAND estimate on the sample of countries for which we have firm data on level of counterfeits. There is again a steady decline from 2006-2008, that now continues through 2009, with the uptick in counterfeiting occurring in 2010 and 2011. The similarity in trends with the RAND model for the smaller set of countries in which we have firm data on counterfeiting is not surprising, but reinforces the point that in general the RAND model is missing the uptick in counterfeiting suggested by firm data in all of the countries for which the product is sold.

As with any model, the RAND model is susceptible to modeling error. However, before drawing a conclusions about the validity of the RAND model, it is important to consider that comparisons made to the firm are based on their estimations as well.[24] We show the impact of this modeling uncertainty in Figure 5-1, which shows the trends generated by the firm data (in red) and the RAND model (in blue), with the 95% confidence intervals surrounding the RAND estimates. It is important to note that the 95% confidence intervals contain the value of the firm estimate in 4 out of the 6 years, despite differences in the composition of products over time. Assuming the firm estimate is the true value of counterfeiting, the RAND model does not perform well due to not capturing the increase in counterfeiting suggested by the firm data in 2008.

---

[24] To the extent that firm estimates are generated through a modeling approach extrapolating from specific markets in which mystery shopping is conducted, it too will have modeling error that is not currently represented here. This is because the firm did not provide us with any information regarding which estimates were model based and the standard deviations of their estimates (i.e. the extent of uncertatinty in the estimates).

Table 5-3: Summary statistics of predicted counterfeits per product from RAND model and firm method

	Mean	Std Dev	N	Year 2006	Year 2007	Year 2008	Year 2009	Year 2010	Year 2011
Panel A									
Firm estimate of counterfeit by product & country (in units)	223,037	492,378		470,891	407,710	340,476	332,896	122,073	57,883
RAND estimate of counterfeit by product & country (In Units)	96,113	187,005		101,692	90,311	133,106	94,891	77,644	120,359
Number of observations (country-products)	464	464		39	47	56	57	196	69
Panel B									
Firm estimate of counterfeit for 5 consistent* products sold in 5 countries (in units)	410,984	733,383		924,866	497,954	419,351	376,807	174,490	62,777
RAND estimate of counterfeit for 5 consistent* products sold in 5 countries (in units)	82,302	129,579		60,211	40,069	29,460	27,561	29,718	38,528
Number of observations (country-products)	127	127		13	21	25	24	25	19
Panel C									
Firm estimate of counterfeit, as a percent of actual units sold	56.4%	20%		84%	55%	64%	54%	37%	50%
RAND estimate of counterfeit, as a percent of actual units sold	41.3%	25%		73%	43%	30%	18%	22%	61%
RAND Full Sample	30%	85%		39%	36%	23%	12%	21%	41%
Number of observations (country-products) - RAND & Firm estimates	464	464		39	47	56	57	196	69
Number of observations (country-products) – RAND Full Sample (FS)	749	749		96	102	118	129	196	108
Panel D									
Firm estimate of counterfeit, as a percent of actual units sold - dropping China	23%	12%		25%	16%	13%	14%	25%	43%

RAND estimate of counterfeit, as a percent of actual units sold – dropping China	35%	19%	47%	34%	33%	18%	19%	56%
RAND Full Sample (FS) – dropping China	28%	96%	32%	29%	28%	11%	19%	43%
Number of observations (country-products) - RAND & Firm estimates – dropping China	404	404	28	36	40	50	190	60
Number of observations (country-products) – RAND Full Sample (FS) – dropping China	705	705	89	95	110	122	190	100

Note: * Consistent refers to the firm obtaining a counterfeit estimate for a product in every year for a country.
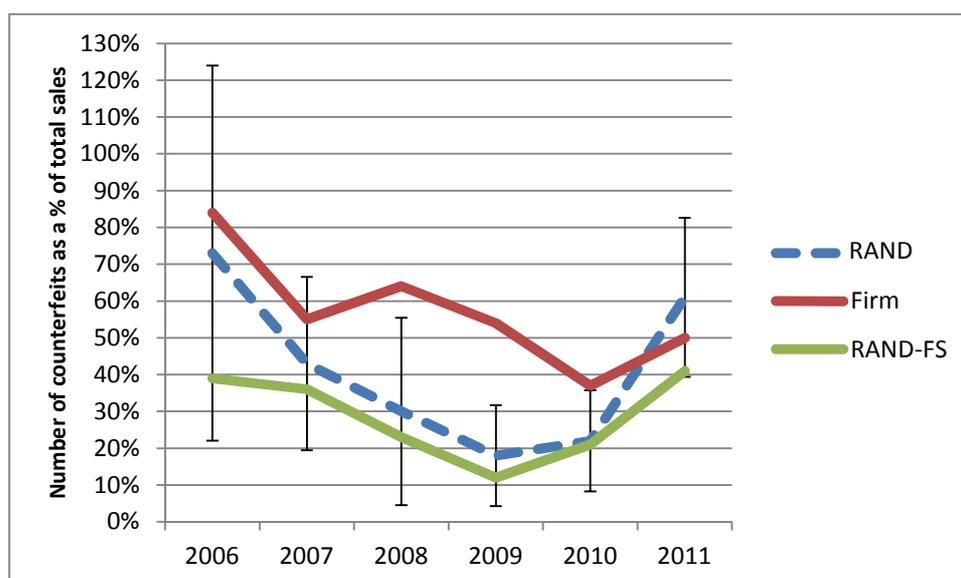
Figure 5-1:   Trends in counterfeit amounts as a percent of actual units sold
Source: RAND model estimate[25] and Firm estimate.[26]
Note: FS = full sample

Sensitivity to infrequent outliers

Given we are comparing to firm estimates, but do not know how these estimates were determined, it is important to understand to what extent the trend reported in Figure 5-1 is accurate even for the firm data. A variety of checks were made on the data, and trends were assessed based on alternative reasonable criteria for outlier deletion. Perhaps not too surprising we discover the firm trend for counterfeiting is highly sensitive to alternative choices for outlier deletions. For example, the firm has products for which the estimated amount of counterfeiting is greater than 5 times the amount of actual sales of the product in a given year. When we remove these from their overall estimate of counterfeiting, the trend for the firm estimates changes considerably (as shown in Figure 5-2 and reported in Panel D of Table 5-3). This simple restriction, which in essence removes a large number of observations from a single country (China), has an important impact on the implied trend to which the RAND model is being held.

Importantly, unlike the trend reported for the firm data, these outliers do not have as large an impact on the RAND model trend, as the model did not perform well originally in terms of estimating these large values in the first place (hence the gross underestimation of levels seen back in Table 5-3). Indeed, the firm's trend without the relatively large estimates now falls within the 95% confidence interval for the RAND estimate in all years, and the RAND model is broadly consistent with the exception of relatively small differences in 2008 and 2009, when the global recession began.

The reasonableness of throwing out these potentially valid observations from China cannot be adequately assessed without knowledge regarding which of these observations were based on mystery shopping data versus the firm's own model extrapolations. This information, however, was never provided to RAND. In any event, the exercise demonstrates the sensitivity of the main

---

[25] Based on modelling based on firm data and publically available data.

[26] Based on an unknown combination of mystery shopping and modelling.

metric to which the RAND model is being validated and suggests that more rigorous testing using data from multiple firms is required before strong conclusions regarding the utility of the model can be drawn.
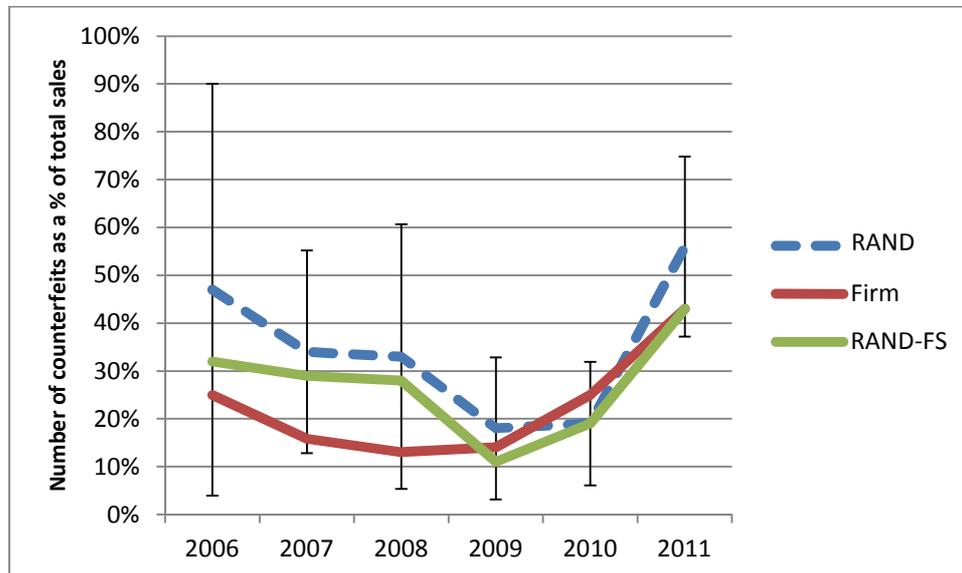
Figure 5-2:   Trends in counterfeit amounts as a percent of actual units sold when China is removed from estimation.
Source: RAND model estimate[27] and Firm estimate.[28]
Note: FS = full sample

5.6   Consideration of an alternative 'simplified' RAND method

Of course, a significant barrier to implementing and testing the RAND methodology as it has been proposed thus far is the data demands required of firms affected by counterfeiting. While many firms regularly engage in forecasting future sales as a means of planning and/or providing information to potential investors (via stock exchanges or by other means), it is highly unlikely that many firms adopt the sophisticated approach demanded for constructing a strong first-stage estimate. Indeed, in our numerous conversations with various industry stakeholders, we were frequently confronted with concerns by the firms about their ability to construct first-stage estimates (unexplained forecast error, adjusted for realized firm and market outcomes). As the idea would be to implement this methodology broadly across numerous firms in various industries, it is important that we assess realistically the data demands of the proposed methodology.

Towards this objective (and to assess the implication of a poorly-specified or non-existent first stage), we evaluated how well the RAND methodology might work if no other information were provided except forecasts and actual units sold. In other words, we skip estimation of the first-stage regression and instead include the difference in the forecasting error as our dependent variable in a second-stage regression.

---

[27] Based on modelling based on firm data and publically available data.

[28] Based on an unknown combination of mystery shopping and modelling.

While we do presume that such an approach would not allow a third party (in this case RAND) to get firm- or even market-specific changes in key variables, it is plausible that unobserved fixed factors influencing specific product markets might still be controlled for using a fixed effects estimation strategy. Thus, we estimate our second-stage model using the same controls as those included in the RAND model, but instead use as our dependent variable the full difference in the forecast and the actual units sold. We also include as additional variables, product-specific fixed effects (as this firm provided us with information on multiple products that are sold to varying degrees in different national markets). The estimated R-squared statistic from this 'simplified' single-equation estimate is 0.13, which might incorrectly be interpreted as evidence that this estimate may have more predictive power than that of the two-stage proposed approach. Such an interpretation would not be correct, however, as the total explained variation from the first- and second-stage regressions using the proposed two-step approach is considerably higher when the two stages are combined. The reduced predictive power of this 'simplified' model can perhaps be better seen in Table 5-4, which replicates results presented in the previous table, now including those for our simplified estimate.

Table 5-4:   Estimated levels of counterfeiting with RAND model and 'simplified' RAND model

	Mean	Std Dev	Year 2006	Year 2008	Year 2010	Year 2011
Firm Estimate of Counterfeit by Product & Country (In Units)	223,037	492,378	470,891	340,476	122,073	57,883
RAND Estimate of Counterfeit by Product & Country (In Units)	96,113	187,005	101,692	133,106	77,644	120,359
RAND Simplified Estimate of Counterfeit by Product & Country (In Units)	186,723	258,935	33,970	203,178	158,259	354,036
Firm Estimate as a Percent of Actual Units Sold	48.5%	0.76	84%	64%	37%	50%
RAND Estimate as a Percent of Actual Units Sold	41.3%	1.07	73%	30%	22%	61%
RAND Simplified Estimate as a Percent of Actual Units	76.2%	0.44	3%	30%	78%	162%

It is clear from looking at Table 5–6 that the RAND simplified model is considerably more volatile, as it is more susceptible to random noise introduced into the data and poorly accounted for by the model. This is due to the inability to control adequately for market factors that might be related to actual sales and spuriously correlated with the counterfeit measures in some years, but not others. The overall level estimate of counterfeit in the market using the simplified estimate is significantly smaller than the firm and RAND approach in 2006 but gets larger than that of both estimates in 2008 going forward. The variation around this point estimate (given by the standard deviation) is also very large, suggesting a very large confidence interval around any point estimate generated by this model.

5.7   In sum

Preliminary evidence from models estimated using data from a single firm operating internationally with multiple products targeted by counterfeiters suggests that the proposed alternative two-step methodology being proposed herein might be a fruitful avenue forward for understanding the size of the counterfeit market, or more precisely: how that market changes over time. It is clear from pilot testing the model with data from one firm for a specific time window, that the RAND model underestimates the firm's own estimates in most years. This appears to be due to a small number of infrequent outliers in the firm data of very high estimates of counterfeiting for particular products in specific markets (in this case China).

Even though the model tends to underestimate the level of counterfeiting in most years, there is some evidence that the model may track the general trends in counterfeiting experienced by a single firm. While the results suggest that the model may miss some important short-term deviations from the broader trend (see: years 2007-2008 and 2009-2010 in Figure 5-1), it's unclear the extent to which these deviations are driven by mystery shopping estimates vs. the firm's own modelling of the data (based on evidence in Figure 5-2 showing no rise in the firm data trend for 2007-2008). In a specification that excluded a single country, China, the RAND model predicts a trend that is very similar to the trend in the firm data. Additional testing of the model will be required using data from other firms before the model's utility can truly be known.

Should the RAND model be found to be reliable in predicting broader trends in counterfeiting across firms, then the model could be extremely useful in its ability to evaluate the impact of policy changes, despite evidence that it could underestimate the total amount of counterfeiting in terms of levels of specific products counterfeited.

While the evidence is incomplete and more piloting is clearly needed, there are a number of benefits associated with using this approach to understand trends, should it be proven robust with additional data from firms. First, the proposed methodology can be implemented at relatively low cost *vis-à-vis* other industry gold standard method of direct sampling of product, such as mystery shopping and chemical product analysis, and thus provides an economically feasible tool for government and regulatory agencies that need to monitor trends in counterfeiting or evaluate the effectiveness of alternative policies and interventions. The data collected and applied for the firm in this study was one-third the cost of mystery shopping data collection and analysis of selective products and markets, and that is simply the initial cost savings. The cost of adding additional products to the RAND model for counterfeit estimation, given the model infrastructure is already in place, would be marginal. Given the modelling approach, additional data would provide further accuracy of the estimation. As such, it is a reasonably cost-effective strategy to estimate counterfeit trends.   For firms, the RAND approach provides an alternative cost-effective means for filling in data gaps in other markets where the gold standard is not applied and does so in a way that is not sensitive to selection issues that can bias estimates using extrapolation methods from gold standard samples. Moreover, not all firms can afford to collect evidence on counterfeiting of their product using gold standard methods because doing them well is so expensive. Thus for those companies that cannot afford sophisticated limited sampling, this methodology provides a cheap alternative that can be consistently applied across markets.

A second benefit of the RAND approach is that it is quite flexible and can be modified to meet unique aspects of specific products, firms or industries while still generating aggregated output that is generalisable across products, firms and industries to generate regional-market or global

estimates of the level of counterfeiting. The flexibility comes about because of the two-stage estimation process. In the first stage, a firm interested in understanding its own deviations from forecasts can customise the information in its first stage to be as firm- or product-specific as desired. The fact that different firms use different explanatory factors to describe why they deviated from their original forecasted output post hoc is irrelevant for estimating the amount of counterfeiting in the second stage, provided that the first-stage model is appropriately specified for that firm. The better firms are at estimating forecasting error, the less noise is generated in the error term that gets pulled from the first stage and used in the second-stage model. If firms are not very good at explaining deviations from expected forecasts, then that just means that the error term in the first stage will have a lot more 'noise' and variance. Provided firms are not systematically biased in their ability to explain forecasting errors (e.g. everyone is good or everyone is bad), then the noise will be random and not introduce any systematic bias in the second stage. Estimates will remain unbiased, although they will have a bit larger confidence interval (error band) due to the additional noise included in some firms' estimates of unexplained forecasting error.

The flexibility is not limited to the first stage, however. To the extent that market factors influencing counterfeiting are unique across industries, the second-stage model could be modified so as to include unique sets of counterfeiting measures for particular industry groups. Aggregation of total counterfeited units sold occurs post estimation of this second stage. The fact that there are different controls on the right hand side of the second-stage model for different industries does not reduce the analyst's ability to aggregate across markets, but merely improves the precision of estimates of counterfeiting (in units) from those specific firms/industries. Empirically, a global model that interacts counterfeiting measures with industry dummies would in essence estimate the model in the same way (allowing for specific counterfeiting measures to influence some products/industries differently than others). Thus, the approach is adaptable to unique market environments and products.

A third benefit of this approach is that it enables a systematic comparison of counterfeiting effects across firms operating within the same market or for similar products. This is because the same model is applied across firms, and hence any general market error that might exist in estimating counterfeiting more generally will not influence the relative effects of counterfeiting on one firm *vis-à-vis* other firms. It is difficult to make direct comparisons of estimates of counterfeits obtained from various firms when unique methods are used by each firm. Differences could be the result of differences in methods rather than true differences in the effects of counterfeiting on authentic good market sales.

There are real concerns about the feasibility of implementing the model on a broad basis, however. And as articulated earlier in this section, there are potentially some real threats to our identification strategy that can only be fully assessed through replication of the model on data from multiple firms and industries. In such an exercise, it will be important to assess the model's relative performance with respect to recognised benchmark methods within different industries. For example, it should be tested for pharmaceuticals, textile goods, technology products and food/drink, to ascertain if the model performs equally well across industries/products. While the specific regression coefficients associated with particular counterfeit variables need not be the same across industries or even across firms (i.e. the flexibility mentioned above), the strategy of making use of a core set of variables to describe the potential risk of counterfeiting across national or regional markets is key. The approach is considerably more useful when a sincere attempt is

made to reduce the observed forecast error by taking out firm- and market-level factors that after the forecasting period will cause the forecast to be inaccurate, as indicated by the very different findings from the two-stage approach and the simplified model shown above. In doing so, the remaining counterfeiting variables are better able to explain variation over time in the remaining unexplained forecast error.

While the proposed method is clearly data intensive, the preliminary evidence suggests that estimates generated from this method could generate meaningful information on movements in the trends for counterfeit goods more generally. Based on data from this firm, a simplified model that ignores firm- and product-relevant factors influencing deviations from forecasts is not as useful for predicting counterfeit levels and does not trend in a fashion that is consistent with either the firm data or the two-stage method. Thus, it appears important that some firm-level explanation of deviations from forecasts is needed for the methodology to generate meaningful estimates. Future work might find the simplified model more useful than evidenced by data from this single firm, but efforts should be spent on the more promising two-stage approach.

CHAPTER 6   **The challenges of measuring IPR infringements**

In this chapter we describe some of the challenges in implementing the proposed methodology. Our description of the challenges is taken from our numerous conversations with representatives of industry stakeholders, conversations with several academic and policy groups who work on issues related to counterfeiting and UUPC and our analyses of the pilot data used in this report. We group the challenges raised by all of these groups into three categories. The first category of challenges is related to the forecast data. The second concerns data sharing by firms. A third category is related to problems that are likely to be unique to specific industries, and we describe some of the more notable concerns raised during our conversations.

6.1   Challenges with using forecast data

At the heart of the proposed methodology is forecast data from firms. If forecast data are not available or if the data quality is poor, then either the methodology will be impossible to implement or implementing the methodology will produce results that are unreliable. In our discussions with firms we found that firms generally had forecast data, but that tracking down the data within the firm was problematic. Forecasts within a firm are potentially produced by different units within the firm. For example two different units in a firm might produced forecasts related to the European and United States markets. Consequently it is unlikely that a single repository of forecasts exists within a given firm. Similarly, historical forecast data are also potentially difficult to identify and not always available, as old forecast data are not always archived and are sometimes destroyed.

Aside from data availability some individuals expressed concern regarding forecast data reliability. If forecasts are not really forecasts and are better thought of as sales targets, then results generated from our methodology are likely to be biased or incorrect. Ideally firms not only provide forecasts but also information regarding the forecast-generating process. Identification of the method for generating forecasts will also likely aid in interpreting both first- and second-stage results from the proposed method.

A final concern relates to the possibility of firms to manipulate their forecast data before submitting them to be included in our model. Ex-ante one might believe that implementation of any forecast-based methodology for estimating counterfeiting will create incentives for firms to manipulate the forecast data in an effort to bias the resulting counterfeiting estimates upwards. Although plausible, there are several potential reasons to believe that manipulation of forecast data by firms might not occur. First, as noted above, implementation of this methodology

requires some knowledge of the forecast-generating process and a reasonable understanding of this could lead to obvious diagnostics regarding whether the forecasts are being manipulated. We expect that the potential for detection should reduce the incentives to manipulate the forecasts. Second, manipulation of the forecast data is not as straightforward as one might believe. In fact manipulation of the data would have to occur in a manner that is directly related to variables included in the second stage of the estimation procedure. If second stage variables are not revealed ex-ante it is not clear that firms will be credibly able to manipulate their data in a manner that generates larger measures of counterfeiting.

6.2 Challenges with obtaining forecast data

Although forecast data seemed to be available in many instances, firms were extremely reluctant to share it. We describe three salient reasons for this reluctance to share data.

First, firms were concerned about the potential for the disclosure of forecasts and sales data. Disclosure of forecasts and quantities sold poses different risks across industries. In nearly all industries competitors or new entrants into the market can potentially use the revealed information strategically to hurt the initial firm. In some specific industries, such as sports broadcasting, forecasts are as important as data on quantities sold since forecasts are one of the main pieces of information used in auctions to determine bids to concession rights. In these latter instances firms will be extremely reluctant to share this information since information disclosure can directly affect the bidding process and lead to a considerable reduction in the bidding price. In two instances firms were willing to share data with RAND upon the execution of a non-disclosure agreement.

Second, as noted above, although available within firms, identifying and understanding the forecast data is likely to be a significant challenge. Identifying the location of the data, obtaining the data, and understanding the data and the data-generating process should be seen as a non-trivial hurdle for firms to overcome. In large firms that process may span many geographic regions: it can and does take up to several months.

Finally, firms seemed to be reluctant to be the first participant in the study from a given industry. In this instance there seemed to be a first-mover problem, as firms were potentially afraid that they might mistakenly be thought of as the only firm in a given industry with a problem of IPR infringements.

6.3 Important miscellaneous challenges

A third set of challenges relates primarily to industry-specific concerns. Addressing these set of challenges directly is complicated and is likely impossible without actually working with the data. Several stakeholders raised the question of predictive power in the second stage. This is an important question and theoretically depends on the sample size and the variables related to counterfeiting that are available for inclusion in the second- stage analyses. If the primary variables that are to be used are macro-level variables, then sample sizes will have to be larger and longer cross-country time series data will be required.

Some unique industry-specific concerns are related to stockpiling of products. If products can be stockpiled for extended periods of time then one would expect that understanding the difference

between forecasts and quantities sold becomes more complicated. We believe that aggregation of data over longer time periods will lessen the extent to which this is a problem. For example one would expect inventory to cause problems if the observation of analysis was occurring over shorter time periods, for example product-months, as opposed to larger time periods such as product-years. Analysts can also consider conducting robustness checks using aggregates over time periods of varying length (quarterly, semi-annual) and lags between forecasts and quantities sold to understand the sensitivity of the results to this concern.

Some stakeholders were concerned that firms take into account counterfeiting in their quantity forecasts. If firms calculate product demand and then reduce it by an amount of counterfeit, this amount needs to be presented, as well as explanation of where it comes from. Even if it is a constant amount or percentage of the total, it is useful for us to understand the previous amount considered and assess whether our model can pick that amount up. We need to know, in particular, whether firms calculated the amount of counterfeiting in terms of demand for authentic goods only or presumed willingness to buy knock-offs. If firms do not consider the amount of counterfeit in the forecast explicitly and simply start, for example, with how much sold previously, this may still be indirectly taking into account counterfeiting and we would need to know this. As noted above, understanding the forecast-generating process is crucial.

A fourth concern relates to parallel trade. Parallel trade refers to the shipment of product from one geographic location to another geographic location where the product is sold at a lower price. We believe that this is a valid concern in specific instances, particularly in cases where unanticipated government levies (for example in markets such as tobacco and alcohol) haven't been taken into account by firms in their forecasts. In cases where price differentials across geographic markets are known in advance and firms explicitly take these differentials into account in their forecasts, parallel trade should be less of an issue. A pilot exercise with data from an industry in which parallel trade is taking place – and preferably compared with an alternative method such as mystery shopping – should assess the extent to which this concern affects the results of our model.

A fifth concern relates to the applicability of the model to UUPC, and in particular online content. In the available literature, the scale and impacts of online UUPC are typically dealt with separately from tangible infringements. This is partly because the theoretical assumptions underpinning markets affected by counterfeiting and by online UUPC are slightly different. Given the prominence of UUPC in the recent debates around IPR infringements and the number of publications addressing this issue, we explore these concerns in more detail than those identified above in the following chapter.

6.4   In sum

While preliminary evidence suggests that the RAND method does a good job at tracking the general trends reported by mystery shopping, a number of challenges remain. These need to be addressed or taken into account before the RAND method can be applied more broadly.

Firstly, the applicability of the method depends on the availability and quality of forecast data from firms. This chapter listed a number of reasons why collecting such data may be difficult. Also, even though this is complex and could lead to detection, there is a possibility of firms manipulating their forecast data before submitting them.

Secondly, although forecast data seemed to be available in many instances, firms were extremely reluctant to share them. Many of the reasons, such as the concern about disclosure of commercially sensitive data, are salient and should be addressed.

Thirdly, some industries have specific characteristics that require serious consideration. Addressing these challenges directly is complicated and will likely require additional data collection. Examples of such industry-specific complexities include parallel trade and stock piling. However, any approach attempting to generate estimates of IPR infringements in a systematic way across multiple firms and industries will have to aggregate results to such a degree that they will be far less precise and meaningful than if the assessment were being done for a single firm or industry.

Finally, given the prominence of Internet piracy in the current debates, it is important to reflect on the applicability of the RAND methodology to UUPC in more detail. While it is too early to dismiss the RAND methodology in this domain, it is important to highlight the concerns and limitations. The following chapter discusses these concerns in more detail.

CHAPTER 7 **Applicability of the model to unauthorised use of protected content**

It is clear from the discussions that we have had with numerous firms concerned about unauthorised use of protected online content that there are theoretical and practical concerns about the applicability of our approach, based on a perception that our underlying theoretical assumptions just do not apply to the online markets for protected content. We believe that the evidence available does not clearly dismiss the appropriateness of our approach. This chapter discusses the theoretical and practical concerns of our methodology related to UUPC. In the absence of any empirical data to test these concerns it is too early to conclude that the RAND model does not apply to markets subject to these types of infringements. Therefore, in addition to identifying these concerns, we suggest a number of recommendations that may be explored in order to arrive at a methodology that addresses the specificities of UUPC.

7.1 Theoretical concerns in measuring UUPC

In this section we discuss the assumptions underpinning our model and their applicability to UUPC by considering the market specificities on the demand-side and the supply-side. Table 7-1 provides a summary of the key elements of our theoretical model that may or may not apply to UUPC.

Table 7-1: Assumptions underpinning the RAND method and their applicability to counterfeiting and UUPC

Economic factor	Description of our model	Sectors affected by counterfeiting	Sectors affected by UUPC
Market structure	• Differentiated competition→oligopoly	✓	
Basis of competition (quantity or price)	• Firms choose a quantity contract and are committed to supply at least that much quantity – price given by their share of demand curve	✓	May compete on price, but that determines quantity (can't pre-determine both)
Firm behaviour	• Low and high degree of heterogeneity (oligopoly and product differentiated competition)	✓	✓
	• Firms consider the total market demand curve for a product and produce a share of total demand	✓	True if creative good is unique (movie, artist song)
Timing	• Two stages: legitimate firms first simultaneously commit themselves to quantity, and afterwards all firms compete on quantity	✓	✓
Consumer behaviour	• A representative consumer purchases products by maximizing utility subject to budget constraints	✓	✓
Market response	• Prices are determined through consumer demand for firms' chosen quantities, although oligopolistic firms have some control over pricing once they determine quantity	✓	✓
Demand schedule	• Downward sloping in its own price	✓	✓
	• Demand increases with increases in the price of competitor illegal firms	✓	✓
Equilibrium	• Authentic firms may maximise profit with counterfeiting levels (UUPC rates) above zero	✓	✓
Determination of counterfeit share	• Share of counterfeits (UUPC rate) is determined by an authentic firm's decisions regarding price, quality and quantity	✓	True in the monetary sub-market. Not necessarily true in the non-monetary sub-market.

7.1.1   Supply-side considerations

When thinking about UUPC, the assumption of some degree of market power by producers is usually justified by the fact that creative products 'are sufficiently horizontally differentiated to make the demand for any particular product largely independent of the prices of other products in the same category' (Belleflamme and Peitz, 2010, p.5). Different degrees of heterogeneity can be observed across firms, where some might compete in very specialised niches while other may serve a variety of tastes, but a certain extent of market power will always be granted by the intellectual uniqueness of a creative product.

One way in which we can determine the market structure is to consider the four- or five- 'firm concentration ratio', which is the percentage of the value of total sales accounted for by the four or five largest firms in an industry. It is a way of measuring the concentration of the market share held by particular suppliers in a market and the rule of thumb is that any industry with over 40 percent market share is considered an oligopoly and less than 40 percent is monopolistic competition (Mahajan, 2006). Counterfeiters have been characterised more often as operating in an oligopoly and monopoly market, with discussions of product-differentiated markets as well. In the music industry for example, one study finds that two of the largest markets in Europe, Germany and Britain can be characterised as oligopolies (Tschmuck, 2012). There is some

evidence to suggest that therefore, at a market level, firms facing UUPC may behave similarly to markets with counterfeiting.

In the legal market, the firm in principle chooses the combination of quantity and price that maximizes its profit, given the total demand curve. It does so facing a demand curve whose downward slope is determined by market power, in absence of which the demand would be perfectly elastic (i.e. horizontal). Because of this market characteristic, a change in the supply by the firm affects the price it can charge for all units supplied. In the illegal market, the firm faces competition by pirates. The price of obtaining an illegal digital copy is not zero: instead, it consists of the opportunity cost of time and transaction costs such as expected fines and hardware threats (i.e. viruses) (Stryszowski and Scorpecci, 2009). Consumers choose between the legal and illegal copy by maximizing their utility subject to a budget constraint whose slope is determined by the respective prices. A firm can thus capture a fraction of the demand destined to the illegal market if it acts on the slope of the consumers' budget constraint by reducing the price (Belleflamme and Peitz, 2010). The profit-maximization problem of the legitimate firm may lead it to decide to choose a price that leaves some of the demand to UUPC.

As a result, the prevalence of UUPC in the market is determined by these economic decisions by consumers. In the case of online UUPC it is worth remembering that 'supply is often driven by factors other than the prospect of monetary profits; (Stryszowski and Scorpecci, 2009) and thus some of the incentives that we embed in our model as driving the entry of counterfeiters might not always be at play for providers of content. However, even though this link is potentially weaker from a theoretical standpoint, it could hold empirically. In fact, the observed firm's economic decisions and the observed rate of unauthorised use could be highly correlated, even if not for strictly economic reasons. For example, it seems reasonable to imagine that the products with largest supply – the most popular – are the most illegally copied.

Part of the literature argues that reproduction of copyright content is essentially the authentic good (Liebowitz, 2005), implying the quality is identical to the authentic product, while other models assume lower quality of the copy (Bae and Choi, 2006). In both cases, some non-zero degree of substitutability remains which triggers the erosion of market shares from the legitimate producers. With substitutes, demand falls for decreases in the price of the competitor; therefore, if the price of pirated material is less than the price for the authentic good then customers would switch (assuming the risk associated with possessing an illegal good outweighs the cost).

However, some models have incorporated features that could have a positive impact on consumers' welfare. For example, it is argued that there are sampling effects with creative goods (Liebowitz, 1985) and that pirated goods behave more like complements, allowing consumers to learn about the creative content and purchase the authentic good. Liebowitz (2005) illustrates a much more complex set of interactions in which sampling leads to a decrease in legitimate sales. Furthermore, Conner and Rumelt (1991) first introduced the idea of networking effects, where the consumer's utility increases with the number of consumers having access to the same product, irrespective of whether such access is provided by legitimate or illegitimate copies. Finally, the seminal work by Liebowitz (1985) also described the idea of 'indirect appropriability', which refers to the higher willingness to pay for the original product by the consumer if he believes he will have the ability to make copies of it. In general, these groups of effects create theoretical conditions under which some non-zero UUPC rate might be beneficial to the right holder.

7.1.2   Demand-side considerations

At the heart of the demand-side aspects of markets with fake goods is asymmetric information. In counterfeiting, the literature tends to focus on two types of asymmetric information: counterfeiters deceiving consumers and counterfeit consumers deceiving other consumers (Qian, 2011). In the case of online UUPC, deception is less likely to play a role when the illegal copy is available for free, but can happen when a purchase is involved from an illegal seller (Stryszowski and Scorpecci, 2009). Furthermore, while taste for the quality of the fake good should in principle show similar variance across consumers of counterfeit and pirated products alike, the cost of obtaining the illicit copy might vary a lot across these two groups. For example, while the cost of willingly purchasing a counterfeit is typically embedded in its price, the cost of an illegal download incorporates time and search skills whose availability varies across consumers, generating different transaction costs (Stryszowski and Scorpecci, 2009).

In terms of observed consumer characteristics, the literature on UUPC has consistently found evidence that 'in most societies the distribution of actors active on the markets for pirated digital products is highly skewed towards young males' (Stryszowski and Scorpecci, 2009). Another important empirical finding is that consumers of digitally pirated goods also tend to purchase more legal copies (Zentner, 2006). These features might not apply to the representative consumer of counterfeit goods, who might look more similar to the average, especially if deceiving purchases happen with some degree of randomness. Even though there might be differences in some consumer characteristics between counterfeiting and UUPC, none of them deviates from the two key assumptions in our model regarding the demand side: (1) consumers choose between the pirated and the original good by maximizing utility and (2) consumers are subject to a budget constraint and thus are sensitive to variations in relative prices.

When seeking to understand whether the aspects of demand are similar between counterfeit goods and UUPC, another avenue is to consider the properties of a public good. A public good is one that is non-rivalrous and non-excludable. Non-rivalry refers to the concept that an individual can consume a good and does not reduce the availability of that good. An example of this is electricity – one person using electricity to light a room does not prevent another person from using electricity to watch television. Non-excludability refers to the concept that an individual cannot be excluded from using the good. For copyright content, such as music, they are non-rivalrous in that one person's enjoyment from listening does not stop another person in the room from being able to listen and enjoy the music. In fact, there are arguments that copyrighted and patented goods have network effects, whereby consumers' enjoyment of the good increases when there are more users of the product (as, for example, with telephones). The literature is unclear as to the effect that individuals' consumption of copyrighted and patented goods have on others' consumption.[29]

Regarding excludability, it can be argued that copyrighted goods are excludable – the copyrighted content can be withheld from an individual. We see this in practice, for example, when producers require payment for the music or software and release dates are scheduled staggered across geographic areas. As compared with counterfeit goods, it can be argued that music and software products are more similar to public goods as they are both non-rivalrous and non-excludable to some degree once they have been purchased. However, to the extent that there are only so many

---

[29] See for example Boldrin and Levine (2003) and Klein *et al.* (2002) for arguments on both sides of this issue.

products produced and sold, the extent to which people might have access could be limited and various protections on reproductions have been developed.

In conclusion, there do not appear to be critical aspects of our theoretical framework that only apply to counterfeiting, although some of the assumptions warrant further consideration using empirical data. First, the firm's commitment to producing a given quantity is less stringent in the case of digital markets where there is almost real-time infinite supply capacity once a master copy of the title (music album, movie, book, broadcasting event) is produced. Nonetheless, firms in the creative industries will also forecast total demand when deciding whether to produce the work of an artist. Second, when the nominal price of the unauthorised digital good is zero, there might be non-monetary factors that we fail to capture theoretically with a simplifying model of economic behaviour. However, these factors might still be captured in our reduced-form model empirically. We therefore conclude that it is worth trying to pilot-test our same approach with empirical data from the creative industries.

7.2 Practical concerns in measuring UUPC

While the theoretical applicability of our proposed methodology to the study of UUPC cannot be excluded a priori, we have not succeeded at piloting it with firm-level data from the companies that we have engaged throughout the project. We have consulted various firms affected by UUPC and their trade associations, as well as experts on the matter. During these discussions a number of concerns were raised in the practical implementation of the RAND model to measure the impact of unauthorised use on sales of protected content. As illegal file-sharing and downloading dominates the debate around UUPC, most of ours revolve around what is commonly known as online piracy. However, some specific comments can be made about illegal streaming of online content as well.[30]

Companies and trade associations in both the music and the movie industry have claimed that their internal forecast process for sales is regarded as highly unreliable and scarcely precise, and often leads to very off-target predictions. According to an expert we interviewed, the lion's share of forecasts in the music industry over-estimate realised sales, because by definition every artist who is signed has an expected positive return but 90 percent of them fail to deliver it. In contrast, the minority of forecasts that under-estimate sales do so because 10 percent of new artists who are signed perform exceptionally well and beyond expectations. Furthermore, the creative industries consider the factors captured by the error term of their forecasts to be extremely difficult to measure empirically. An example often made is that of quality, which can ex-post explain a lot of the discrepancies between forecasts and actual sales. We have reiterated that those ex-post observables are exactly the types of factors that we aim to consider in our first stage in order to compute an unexplained forecast error. However, it appears that quality and other such variables

---

[30] Our interaction with a right holder in the sector of sports events has led to the identification some difficulties in the application of our proposed methodology to online streaming and recommendations to overcome them. The right holder does not produce forecasts, but sells the broadcasting rights to television channels across countries for a defined period. Empirical data are available for the value of those multi-year concession contracts for broadcasting rights over time and across countries. Provided there are enough observations it may be possible to apply an approach based on the same principles as described in Chapter 5, using data for the value of these concession contracts. Pilot-testing would have to assess the validity of this approach.

are not operationalised by companies in a way that could feed into our model. We received similar indications about the nature of forecasting by the movie industry representatives we interviewed. Our understanding is that there is a relevant role played by chance, analogous to what was reported for the music industry, so that the vast majority of movies yield negative profits. However, it is important to note that if a significant part of the large variation in the forecast error (the difference between forecasts and actuals) is 'random' (i.e. unrelated to the observable factors we propose to control for), then our second stage model would be able to explain less of such variation and thus correctly attribute less of it to UUPC. It is unnecessary but probably useful to reiterate that our goal would not be to maximize the amount of variation explained by UUPC, but to get a consistent estimate.

The second type of concern expressed by companies refers to their own estimates of UUPC. In particular, according to the experts we interviewed in the music industry, UUPC is already factored into the forecasts but not explicitly quantified, so it would not be possible for us to account for it. In our dialogue with the companies, we have highlighted that we would need to know if and how they adjusted their forecasts to take UUPC into account ex-ante. However, the position of the company in the music industry that we interacted with is that their assessment of UUPC is built into their forecasts as an underlying factor. The interpretation we provided was that such factor could be thought of as a trend from which our model would be able to identify deviations, which would still be informative from a policy perspective. However, the representatives of the particular music industry stakeholder believed this would not be an appropriate use of their forecasts.

Finally, stakeholders in both the music and the movie industries have highlighted the need for our model to select appropriate second-stage variables that account for the respective sector specificities. Again, in principle there is no technical constraint on our model that would prevent it from being adaptable to these peculiarities. One candidate for exclusion from a model of UUPC is the variable that measures the burden of customs' procedures (see Section 5.3). We know that, with the deep transformations that have occurred in these industries, the importance of online infringements is now almost exclusive with respect to the use of physical supports, so that the movements of infringing content across borders is of lower practical importance. Thus, in the case of UUPC our model could rely on different measures related to the supply of infringing content (in addition to broadband penetration, which is already included in the baseline version our model presented in Chapter 5). As a possible starting point, we could experiment with other variables that have been shown to have some correlation with software piracy, such as measures of R&D intensity, education and bilateral investment treaties (Olsen, 2005, Table 3). Of course, the statistical performance of any variable can be assessed only through actual analyses, therefore while we are aware of these important considerations, we have not been able to address them beyond this abstract level.

The extensive dialogue we have been conducting with the music and movie industries did not generate an opportunity to test our methodology with empirical data, and for this reason in what follows we outline what the next steps would be for exploring alternative ways. The considerations when thinking about generating of an estimate for these sectors are: (1) methods that guarantee the highest standards of statistical consistency, (2) the ability to be repeated over time so as to track changes in trends over time, and (3) the ability to cover all EU countries.

7.3   Recommendations to extend existing methods to UUPC

Based on our extensive reviews of the academic, non-academic and grey literature, and on our conversations with experts, trade associations and firms in the music and movie industries, we have made an effort to identify potential solutions to the key shortcomings of existing methodologies currently applied. This is as an alternative route for improving future efforts to estimate the size of UUPC in lieu of our proposed methodology, which remains our preferred option. In the following section, we describe the type of theoretical and empirical work that would need to be conducted to adjust the existing methods.

The existing methodologies currently applied by the industry essentially consist of a simple arithmetical operation which multiplies the total number of infringements by the number of lost sales per infringement (the so called 'substitution rate') in order to derive the overall impact of illegally downloaded files on sales. This is, in a nutshell, the methodology that was applied, for example, in the study by TERA Consulting (2010). This type of estimation exercise is well regarded by the industry, but as it currently stands presents a number of critical issues that pose significant threats to its reliability. An extensive treatment of these methodologies and their limitations can be found in Appendix A.

The attention shall be drawn here on two main aspects: the number of infringements and the substitution rate.

Number of infringements
The number of infringements is a key input and we need a consistent way of measuring this over time and across countries. There has to be a structured discussion between the industry, policymakers and consumers about the quality of these measurements, which require maximum transparency and methodological rigour. There is a need to close the gap that is left wide open by the scarce information that is typically provided in the studies commissioned by the industry.

Technology has to be a defining driver of this measurement effort. Because unauthorized access to protected content in these sectors increasingly happens by means of digital appropriation, and as technology exists to monitor online behaviour of consumers, the gold standard should be modelled around these capabilities for data capture. For example, the music industry contracts market research companies to monitor illegal file-sharing over the Internet. In particular, software can be installed on the consumers' equipment to record all activities in the background. While we acknowledge the existence of comprehensive panels that guarantee coverage of the entire EU market in virtually real time, there are two aspects that will require particular attention. First, there is a concern that consumer behaviour is modified by monitoring, because this type of data collection happens conditional on the user's consent. Second, the sampling design of the data collection needs to go beyond convenience sampling in order to guarantee the capability for statistical inference. An example of a possible data provider is Nielsen. Nielsen collects data that allows one to study patterns of infringement over time by virtue of following the same sample. Records of visited URLs and time spent on each of them provide a basis not only for counting the number of infringing accesses but also to assess their intensity. Google Analytics, on the other end, can provide insights into the behaviour of the marginal consumers (those who seek orientation in finding a given illegal download as presumably they are new to it) by monitoring searches.

Reliable estimates of the size of illegal downloading are a necessary condition to pursue this methodological route. It is vital that comparable data is collected and organised. However, as we concluded in our chapter reviewing methods, given that data tend to be incomplete and inconsistent over time it is also valuable to triangulate information from different sources.

Substitution rate

The substitution rate also needs to be estimated in a consistent way across countries and over time. It is clearly not a constant, but depends on consumers' preferences. In principle, preferences vary at the individual level, and it is not clear what the constant measures used in many of the studies we reviewed capture about the distribution of preferences (Is it the average? Is it the median?). In order to synthesise this wealth of unobservable information we could aim at capturing its variation across a number of key dimensions. For example, we know that taste for specific styles of music varies by age group – e.g. as consumers age, they are likely to appreciate legal music more because they might attach a higher value to quality, they are less budget-constrained and they might have access to better means for reproducing it (for example a better hi-fi system). Furthermore, taste varies by country due to cultural background, respect for authority, support of the music and arts, and other market characteristics.

It would thus be necessary to focus research efforts on the estimation of these substitution rates. This is especially important given that in the academic literature no consensus has been reached on measurement. As Dejean (2008) explains in a comprehensive review of the empirical literature on piracy, there are issues of comparability due to different data sources and of scientific validity of the methodologies, which have led studies to under-estimate or over-estimate the impact of piracy on sales for a variety of reasons.

At the micro-level, substitution rates can be measured through primary data-collection efforts aimed at eliciting revealed preferences from consumers. This has to be done in experimental settings in order to guarantee internal validity. The RAND model explained in Chapter 5 is an example of this type of study. Random sampling from within each country is necessary for proper representation of potential differences across countries. Equal attention should be paid to external validity, so that the results are generalisable to the entire population. The limitation of exercises of this type resides in the potentially high cost of conducting them across all countries and on sufficiently granular strata of the population to capture the variation described above. Furthermore, they have a high cost of replication over time and of sampling beyond convenience. Studies conducted on a sample of graduate business students that take the professor's class are cheap for recruiting but hardly informative about the rest of the population. However, sampling over time does not have to occur annually, particularly if tastes do not change systematically in short periods of time. It may be that studies every five years are adequate to capture important shifts due to ageing of cohorts and/or technological advances.

At the aggregate level, econometric studies of the causal effect of illegal file-sharing on sales can provide information on the so-called displacement effect, which measures the number of sales lost due to an illegal download. They have the advantage of a relatively low marginal cost of replication once their methodological framework has been fully developed and tested. A number of these studies can be found in the academic literature,[31] where a heated debate has been

---

[31] See for example the work by Liebowitz (2008) and Waldfogel (2010), who both provide an overview of the literature in addition to their own results. It is not our aim to provide a comprehensive review of the literature here.

continuing for a few years over the consistency[32] of the estimates. Almost all studies find a negative effect of file-sharing on sales, with one exception (Oberholzer-Gee and Strumpf, 2007) which finds no effect but has been heavily criticized (Liebowitz, 2010). Using econometric techniques can contribute to closing the knowledge gap about substitution rates, but a simple regression of sales on a measure of infringement will not yield a consistent estimate. In its simplest terms, the econometric problem is one of endogeneity, which happens when a third, unobserved, factor is related to both the dependent and the independent variable and draws a spurious correlation between them. For example, in the specific case of music piracy, it is likely that heavy illegal downloaders also consume a lot of legal music on average, which would lead us to find a positive correlation between illegal downloads and legal purchases and erroneously conclude that UUPC has a positive effect on sales. Econometric techniques can be used to deal with this empirical problem. For example, some studies of music piracy employed instrumental variables techniques.[33] While it looks unlikely that the academic community will develop consensus around an empirical strategy in the short term, there would be value in applying different approaches to the best available data in order to compare their performances in a fully transparent way. Subject to the availability of data, econometric models can estimate country-specific parameters and even stratify the analyses on other variables of interest (such as demographics), which can then be related to country- and group-specific infringement levels to arrive at an estimate of the effect of UUPC on sales.

To mitigate potential biases, uncertainty should be taken into account by incorporating a range of estimates instead of a point value for the substitution rate. Furthermore, the sensitivity of results to different assumptions about the consistency of the econometric estimates should be tested by introducing parameters that weight the displacement rate. For example, if we believe that a given proposed estimator measures a larger effect than the true one, its estimate can be multiplied by a factor between 0 and 1.

7.4   In sum

In this chapter we examined the peculiar characteristics of the industries affected by UUPC and discussed ways in which our proposed methodology might need to be adapted in order to be consistent with them. On theoretical grounds, we do not reject the possibility that our methodology might offer sensible insights on the extent of IPR infringements in these industries as well. However, from an empirical standpoint, while we have received some input on how to tailor our model, we have not been able to test it with actual data. Therefore, even though we tried to provide recommendations on how to improve other existing methodologies, our preferred approach remains our proposed methodology, which cannot be dismissed without pilots. In the

---

[32] Consistency is an asymptotic concept in statistics which describes the ability of an estimator to converge to the true population parameter. In common language it is often mistakenly referred to as un-biasedness.

[33] In order to mute the effect of third, unobserved factors that are related to both the dependent and the independent variables, instrumental variables (IV) can be used. An IV is a variable that is related to the dependent variable only through the independent variable, but not directly. For example, Rob and Waldfogel (2006) use access to broadband as an instrument for downloading. This allows using only the variation in the independent variable that is 'triggered' by the IV in order to explain the variation in the dependent variable, in this way avoiding the endogeneity problem described above.

next chapter, we provide an ideal description of such pilots (Section 8.2) and further develop the discussion of the explanatory variables for the second stage (Section 8.3).

CHAPTER 8 **Next steps towards measuring IPR infringements**

As discussed in Chapter 4, a variety of methods have emerged to attempt to estimate the size of the markets for specific IPR infringements as well as the harms generated from them. As is true for any illegal market where direct measurement is hindered by the lack of legitimate markets, the simultaneous development of alternative promising approaches should be encouraged rather than discouraged, as it is only through a triangulation of information gleaned from each of these approaches that a more confident assessment emerges. While each of the alternative approaches has clear limitations, weaknesses and biases, they can together provide a more reliable understanding of the market than any one of them could provide on its own.

In Chapter 5 we offered a new methodological approach that draws solidly from economic theory of firm behaviour. The core advantage of the proposed methodology is its ability to construct aggregate measures of counterfeiting across multiple products, firms and industries in a reasonably low-cost manner that can be systematically replicated annually in order to gain a better understanding of trends and changes in the IPR market over time. The preliminary assessment of the empirical model derived from this theory was substantially hampered because we were able to complete a pilot-test with only one industry partner. It is our sincere hope that the summary of our work presented in Chapter 5 combined with a more technical derivation and examination of the approach (provided in Appendix A) will stimulate discussion and encourage other industry leaders to consider pilot-testing the method themselves. A true ability to demonstrate the feasibility and reasonableness of this approach will be achieved only if the methodology is tested and demonstrated across multiple firms and industries. We provide in Appendix C some instruction on how a centralised organisation or agency like the Office for Harmonization in the Internal Market (OHIM) might be able to pursue a more intensive pilot evaluation of the model for multiple firms and industries. For a successful pilot to occur, we believe that a number of important next steps must be taken in the coming months or years, and we lay out these next steps here.

8.1 Build trust and buy-in from the industry

While industry experts are keenly aware of the significant limitations of all previous methodologies for measuring the problem, they are also highly suspicious that any new method will generate more promising results. Indeed, there appears to be a very clear conflict between a desire to show policymakers the importance of the problem across industries and firms and the desire to demonstrate how unique the problem is for their own business. Any approach

81

attempting to generate estimates of IPR violations in a systematic way across multiple firms and industries is going to have to aggregate measurement issues to a level that will be far less precise and meaningful than if the assessments were being done for a single firm or industry. That being said, it remains completely unknown whether the higher level of aggregations will necessarily generate biased estimates of overall quantities or general trends. The following thought experiment demonstrates this point fairly well. Think about trying to put a single price per share on the value of Coca Cola Company or British Motors. Each of these companies sell a multitude of products in a variety of markets and the value of their products is influenced by a multitude of factors, many of which are not the same across products. But does that mean that a valuation of the price per share of the company could not be obtained in a comparable fashion for each of these companies? The fact that these companies have both been sold on national stock exchanges says the answer is clearly "No". So too we believe it is possible to develop a methodology that allows individual firms to consider what sort of unique aspects influence their own susceptibility to variable forecasting errors and aggregate up across firms and industry broad macro-factors that systematically influence all markets, such as IPR violators.

Thus, a critical next step necessary to make any further progress in developing the methodology is to build trust and create buy-in from key industry leaders. One way of achieving this is through academic and policy meetings where experts knowledgeable about the problem, the data and the measurement issues actively discuss and debate the proposed methodology and its strengths and weaknesses. To do this, the method must be made explicit so that it can be replicated in a systematic way, which is what we have attempted to do for the academic audience in our NBER Working Paper (Pacula *et al.,* 2012). By making the proposed methodology publically available in a prominent forum where other significant academic work has been conducted on the topic, we are actively engaging academic leaders to provide their perspectives on the approach. The fact that there are weaknesses in the approach is, by itself, not a limiting reason to stop further exploration of the method. Indeed, many valuable models have evolved over time as additional input and insights are gleaned through the practical implementation of the method.

The development of the approach requires more than just an academic discussion of the model and issues. For empirical development to occur, there must be a willingness on the part of industry to test the method, either themselves or with academic partners. The utility of the method for firms and policymakers can be understood only through its empirical testing using real-world, sensitive data from multiple firms. Some steps have already been taken to try to engage and encourage the private sector to test the model with their own data, including a publication with a non-technical explanation of the IPR model in intellectual property media (Schneider, 2012). This sort of direct communication followed by presentations to the policy community and stakeholders may help generate sufficient interest from relevant industry partners that a more rigorous testing and further development of the model occurs.

Finally, it would be advisable for the European Commission and/or OHIM to solicit constructive feedback from industry stakeholders, who –at the end of the day – know their markets better than anyone else. The Commission and/or OHIM may organise dedicated workshops focusing on the applicability of the methodology in specific (clusters of) industries. These clusters may be defined as narrowly as necessary, but should include at least:

- Apparel, Footwear and Designer Products (Textiles);
- Automotive;

- Consumer Electronics;
- Food and Beverage; and
- Pharmaceutical products
- Creative industries (film, music, etc.)
- Other industries

The first five clusters contain products sold by 18 of the top 20 branded values in 2006 according to data from Interbrand and are industries well understood to be targeted, based on seizure data.[34] The sixth cluster should be included on the basis of the distinct market characteristics of Internet piracy.[35] In these workshops, the RAND methodology can be discussed in detail and benchmarked against alternatives, on criteria such as: costs, data requirements, labour intensiveness, complexity (input criteria), reliability, replicability, comparability, and sector specificity (output criteria). These workshops should lead to the selection and execution of additional pilot tests in each cluster, and eventually to the adoption or adaptation of the model according to sector specificities.

8.2   Continued development of the methodology

The pilot RAND Europe conducted of the methodology was useful for demonstrating the feasibility of the approach. However, it was substantially limited by the participation of only a single industry leader. It was therefore not possible to assess whether a core set of aggregate counterfeit indicators can be identified that systematically describe susceptibility of markets to counterfeiting across industries and/or products. Thus, it will be important for the Commission to consider extending the pilot work conducted here in the near future with a few select firms to further confirm proof of concept. To this end, we provide in Appendix C a detailed description of how to go about initiating such a pilot in terms of firm selection and modelling. Our recommendations would be to focus initially on including industry leaders that represent first-movers, so that the modelling assumptions are clearly tested and greater initial attention can be given to the identification of a core set of counterfeit variables for the second-stage model. The industry leaders should represent a broad range of products, however, in order to more fully explore the reliability of using a common second-stage model to describe disparate products. Assuming that a core set of variables is found to be consistently useful for predicting unexplained forecasting error, then efforts can be broadened to assess the reliability of the approach in industries with lower concentration ratios (i.e. more competitive markets).

Of course, even in an expanded pilot a number of important methodological issues will emerge that can be more carefully considered than what we were able to do with only a single firm. In particular, issues may arise regarding the standardisation of product units across different product types, as forecasts might be done on factors other than product units themselves (e.g. euro sales,

---

[34] See http://www.interbrand.com

[35] It should be acknowledged that this distinction is not exhaustive and excludes a large number of important industries (e.g. software, crop protection, computer hardware, power tools, toys, etc.). Depending on the resources and time available, and the number of firms interested, additional clusters could be included separately or as an "other" category. The product areas identified in CEBR (2002) and the OECD sectors with the highest general trade-related index of counterfeiting and piracy of products (GTRIC-p) as identified .in OECD (2009) can be used as general guidelines to selecting and clustering these industries.

where goods can be bundled at alternative prices). Understanding the extent to which alternative ways of capturing forecasts and forecast errors is something that may or may not prove methodologically challenging. Similarly, the extent to which all industries rely on and produce annual forecasts is not entirely clear. Thus, understanding the extent to which the concept of 'unmet demand' can be operationalised will be an important factor to consider in the next stage of development.

By extending a pilot to multiple product groups and industries, it will also be possible to consider the extent to which unique industry characteristics (e.g. stockpiling of products, long supply chains) might impede the implementation of this approach. Most of the product-specific issues that were raised with us by industry leaders we attempted to engage in our original pilot were factors that could easily be dealt with statistically in the model without additional data collection from the firms. This is because statistical models can easily account for unique factors that are time-persistent by product line or firm using a statistical tool called fixed effects. Precise measurement of these factors is not required, provided they are unique to the firm/product and are time-invariant. Indeed, such unique issues were identified across the limited range of products we assessed in our own small pilot, and, as was discussed in Chapter 5, the unique product-specific fixed effects were indeed statistically significant in both our first-stage model and the simplified model. So, some of these factors may be easily worked around through statistical modelling rather than complicated data-gathering tasks.

It certainly may be the case that the methodology developed does not sufficiently address the issues or concerns related to online UUPC. While in theory we believe that the core underpinnings of the model still apply, as described in Chapter 5, the empirical implementation of the methodology in terms of derivation of forecasts and factors influencing them may not be completely transferable. Otherwise, it may depend on the type of online content considered. Without an attempt to apply the general methodology to a few potential goods (movies, music or software) and assessing estimates obtained from these methods to those using gold-standards from within the industry, we will not be able to understand to what extent the above-mentioned concerns are indeed legitimate. Thus, in our mind, an important next step in the development of the methodology is a careful piloting of the approach in a few select markets for online content provided by industry leaders.

8.3 Possibility of tailoring model to sector specificities

As has been mentioned previously, a key strength of the RAND methodology proposed is its flexibility to handle both unique industry/firm factors and market-level factors. While much of the discussion so far has been on the identification of common aggregate measures of counterfeiting that can be used to describe a national-market susceptibility to IPR violations, there is no reason why the RAND model could not be estimated on a sector-by-sector (or industry-by-industry) basis rather than in the total aggregate. Indeed, such a formulation of the model would likely be more appealing to many industry leaders as it enables the model more explicitly to consider some of the unique sector attributes that make specific vulnerabilities more relevant than others (e.g. broadband usage for pirated goods). As the goal of the second stage is to build an estimate of the predicted counterfeited units likely to be available in national markets for various products, what is required for total market aggregation is simply the measurement of counterfeiting in units of product, not the specific controls included in the right hand side of the

equation. Increased precision will come from the second-stage model if we have multiple products being evaluated in terms of the same set of controls (as the increased number of products in countries over time increases the statistical power of the model). However, sufficient power may be obtained even if the textile industry uses a different formulation of the second-stage regression than the pharmaceutical industry or the food and beverage industry. Assuming a large global spread of products (to increase variability in the core counterfeiting measures) and sufficient observations over time, the estimation of the model by sector will likely be a very useful strategy for implementing the model.

To test the feasibility of a sector-specific second-stage model instead of a single total-market specification, in Appendix C we recommend designing a pilot that by construction includes firm-level data from several clusters of sectors based on similarity in the structure of market demand, the types of products sold and/or the supply chain. Of course, other relevant sector-specific characteristics might be considered, but the main objective is to identify a sufficient number of firm products within each sector to facilitate sector-specific models. To the extent that the firms chosen have their own preferred strategies for estimating the effect of counterfeiting on their own bottom line, then it will be possible to compare findings from these sector-specific models with estimates from these other approaches. Furthermore, the models could also be compared in terms of predicted values to a more aggregate global version of the second stage to see if a smaller set of broad indicators (like those used in our analysis shown in Chapter 5) reduces the power and performance of the model substantially.

Multiple products from within each cluster should be used so that clustered sector-specific effects can be uniquely evaluated independently from product-specific effects. Of course, an important question is how many products are necessary to include in order to separate these effects. The answer is not clear and will depend on a number of factors including the geographic variability and temporal variability of the product data. To the extent that firm-specific product information is available over longer periods of time and for a larger number of geographic areas, then fewer products may be needed. However, it is usually the case that a single firm offers several products within a sector (as was the case in our own pilot), thereby enabling estimation with just a few firms providing data.

8.4   Facilitating data delivery

In addition to further testing of the proof of concept, improvement of the general methodology and tailoring it to sector specificities, a useful focus of future work is the consideration of a process in which the data necessary to implement this model can be consistently collected and reported to an agency responsible for implementing the model on behalf of the European Commission. The process involved in identifying the data required for this pilot, collating them in the correct format from the firm and properly structuring them for estimation in the model has been relatively time-consuming and cumbersome as it is was not immediately clear to the firm that it had what we needed to estimate the model. A number of factors contributed to this, including: (1) a lack of understanding by the firm of the data needed and the extent to which the data the firm was able to provide could in fact be used; (2) the fact that forecasts get updated throughout the fiscal year, due to which there were concerns regarding which one was 'right'; (3) concerns that counterfeiting might already be included in the forecasts generated; and (4) an understanding of why data measured at a national rather than regional level were needed. It has

taken considerable time and resources to discuss with the firm (and other potential partners) what the data requirements are and that they could be delivered in relatively flexible formats. Once the indicators, their units, and time series were understood and agreed upon, it took time for the researchers to understand the firm's considerations when building a forecast (and hence what might drive predicted forecasting error in the first stage). Our experience doing this exercise with a single firm suggests to us a process that might facilitate and accelerate data-gathering in the future.

First, it will be important for at least one research team member to spend time with each firm initially to engage the firm, understand its forecasts and sales trends (specifically, the factors that the firm believes influence them) and how data describing those trends are captured by the firm. This requires staff with the capabilities and skills to explain the methodology in a language that is understood by the firm, once a firm 'context' is understood by the staff. If the staff are able to understand the firm 'context' before pushing the empirical model, then the identification and acquisition of data will probably be an easier process.

Second, robust provisions, including signed data-use agreements, are required for data protection. As mentioned above, the model does require some highly sensitive commercial data from industry stakeholders. These firms will be reluctant to share any of such information without an extremely robust system in place that ensures that these data are accessible only to those who are authorised. For the same reason, the system requires a standardised provision for a non-disclosure agreement. An institution such as OHIM would be well equipped and experienced to facilitate these provisions.

Third, once multiple firms are brought online, a standardised template for data submission to a central agency, such as OHIM, should be prepared to facilitate the delivery of data in a systematic way across all firms. We have prepared a data-request questionnaire that might be useful as a starting point for eliciting information about the firm in terms of how it reports units sold, its forecasting methodology and information used to develop it (included in Appendix C). It has been used with a number of potential pilot firms to initiate discussion. Based on what firms report on this form in terms of their use of forecasting, one of three different types of information gathering tools might be used for collecting the actual data for analysis.

- The first format would be given to firms which have their own sophisticated method for forecasting and studying forecasts, in which case firms would simply report to OHIM their own, self-evaluated unexpected forecasting error that already excludes sales that could be explained by post-period realisation of firm-, industry-, and broader-market-shocks, but includes potential counterfeit sales. These are the ideal data for our model, but it is unclear to what extent key firms targeted by counterfeiters fall into this category.

- A second format would apply to those firms which do not have the resources to evaluate forecasts internally but still engage in sophisticated forecasting activities, in which case more specific information about what the firm considers in building its forecast, what unexpected shocks occurred post their forecasting period, and the extent to which counterfeiting might already be reflected in their forecast would need to be gleaned as part of the information-gathering process. This was the approach taken with our industry partner in our own pilot. While this approach required some labour time to understand the business and collect relevant data internally from the firm based on several conversations, it was feasible and

successful. We, as outside analysts, estimated the first-stage regression ourselves, guided by knowledge and input from the firm.

- A third format would apply to those firms which do not systematically develop forecasts, or merely build them based on performance in the previous year(s). For these firms, simple information on differences between units sold and expected units sold (based on either last year's sales or planned shipments to a market area) could be gleaned and a version of the simplified model presented in Chapter 5, which would include firm fixed effects to capture time-persistent unobserved biases in the reporting of information, could be estimated. Again, it would be necessary for the outside analyst to communicate with the firm in order to understand to what extent the firm can represent, in terms of units sold or not sold, unexpected events that arose during the forecasting period so that inclusion of these factors can be considered in a modelling framework. Information from other firms within the same sector may assist in identifying industry- or market-shocks that are also relevant for understanding unexpected growth or declines.

While the model will clearly perform better and the approach will be more promising if more firms targeted by counterfeiters are able to report data in one of the first two formats, it is unclear to what extent this will be the case. However, it may be more important that certain industry leaders report data in the first two formats than that all firms do, as such industry leaders are those which are most likely to be aggressively targeted by counterfeiters. As all firms are not equally targeted by counterfeiters, it will be more important for understanding the size of the market that the model can do reasonably well estimating trends from those that are targeted most aggressively.

8.5 In sum

An English cleric, Charles Caleb Colton, is credited with saying, 'Imitation is the sincerest form of flattery' over two centuries ago. While the persistence of this phrase suggests an element of truth, it is also the case that when the object of imitiation is protected by IPRs, such imitation is not welcomed or legally allowed.

In this report we develop and test a promising methodology for estimating trends in the size of counterfeit markets over time, but the work is far from complete. There are legitimate questions and concerns about the feasibility of the methodology being implemented broadly for estimating trends in a regional market, such as the European internal market, that can only be addressed through a broader piloting of the work. In this chapter, we attempt to lay out what we believe are important next steps that should be undertaken by a centralised organisation with strong industry contacts, like OHIM. They include:

- obtaining greater industry buy-in and participation in a more comprehensive pilot (particularly of key firms heavily targeted by counterfeiters);

- construction of data instruments that can collect information systematically from participating firms in a common format;

- protocols and security measures to protect firm-sensitive data that is shared;

- continued refinement of the methodology and analytic testing of the model, including evaluation of modifications allowing for industry-specific factors to be appropriately considered.

These are not small tasks, but given the magnitude of the issue and the potential harms stemming from the unlawful replication of many goods, the efforts are not unwarranted. The evidence from our work suggests that the pay-off may be quite substantial in terms of development of a new methodology that enables a systematic comparison of counterfeiting effects across firms and industries, thereby providing a better understanding of the overall magnitude of the problem.

There are obvious weaknesses of the approach developed thus far. For example, the RAND model appears to underestimate the size of the counterfeit market in most years (when outliers are not removed). Further testing of the model may not change that result. It is, however, extremely difficult in black-markets to measure accurately the size of the market, regardless of the good. Thus, it is not clear that firms and policymakers should care what the size of the market is in a given year more than whether the problem is getting worse or better. Because of its use of generalisable output units across firms and industries and the systematic approach that can be replicated year after year and country after country, the RAND model appears to be a useful and cost-effective tool to estimate changes in counterfeiting. This type of knowledge is invaluable to firms and policymakers that are eager to understand if the specific policies, systems, or interventions they put in place are effective at shrinking the level of counterfeiting and hence worthy of further investment. Therefore, we believe that the real contribution of this work and the reason why further development is warranted is its ability to assist firms and governments in their efforts to understand what works in terms of managing the problem of illegal reproduction and sale of goods in a dynamic and integrated world market.

The RAND model broadly follows the trend identified by a combination of firm mystery shopping and modelling, but misses some year to year deviations, particularly 2007-2008 and 2009-2010. It is particularly sensitive to what appear to be outliers in the data. When the RAND model utilises data without potential outliers and tests against firm findings, the resulting trends appear to be consistent with trends in the amount of counterfeiting produced by the firm. It should be noted that the model is developed with the intention to aggregate results across multiple firms within the same market. The aggregated results will be much less sensitive to individual firm outliers than in the case of estimates for a single firm.

Moreover, should further testing of the model reinforce the model's ability to track trends over time, this model provides a far more cost-effective way of understanding trends in the counterfeit market than existing industry gold standards. When coupled with periodic information obtained from industry gold standards implemented in particular markets over time, the nature of the systematic bias associated with this RAND model may be identifiable, which would make it possible to scale-up estimates from it and better approximate the size of the counterfeit market in those industries where additional data are available. Doing so will improve the model's ability to achieve the initial objective of understanding the size of the market at a point in time.

REFERENCES

Reference list

Al-Jabri, I. M., and A. H. Abdul-Gader, 'Software Copyright Infringements: an Exploratory Study of the Effects of Individual and Peer Beliefs', *Omega,* Vol. 25, No. 3, 1997, pp.335–44.

Allen Consulting Group, *Counterfeiting of Toys, Business Software and Computer and Video Games*, Report to the Australian Toy Association, the Business Software Association of Australia and the Interactive Entertainment Association of Australia, Sydney, 2003.

Attributor, *Demand for Pirated E-Books Research*, Industry Report, Attributor, 2010.

Bae, S.H., and J.P. Choi, 'A Model of Piracy', *Information Economics and Policy,* Vol. 18, No. 3, 2006, pp.303–20.

Bale H., *Pharmaceutical Counterfeiting: Issues, Trends, Measurement*, WIPO/OECD Workshop, 2005. As of 9 March 2011:
http://www.oecd.org/industry/industryandglobalisation/35650404.pdf

Bate, R., 'Making a Killing: The Deadly Implications of the Counterfeit Drug Trade', Washington DC: AEI Press, 2008.

Belleflamme, P. and M. Peitz, 'Digital Piracy: Theory', SSRN eLibrary, 2010. As of 9 March 2011:
http://ssrn.com/paper=1698618

Bhattacharjee, S., R. Gopal, K. Lertwachara, J. Marsden, R. Telang, 'The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts', *Management Science*, Vol. 53, No. 10, 2006, pp.1359-74.

Boldrin, M., and D. Levine, 'The Case Against Intellectual Property', CEPR Discussion Paper No. 3273, 2002. As of 9 Mar 2011: http://ssrn.com/abstract=307859

Bryce, J., *Deliverable D3: State of the Art Report*, Counter, 2009.

Business Software Alliance/IDC, *Seventh Annual BSA/IDC Global Software Piracy Study*, May 2010.

Burke, A.E., 'How Effective Are International Copyright Conventions in the Music Industry?', *Journal of Cultural Economics,* Vol. 20, No. 1, 1996, pp.51–66.

CEBR – see Centre for Economics and Business Research.

Centre for Economics and Business Research, *Counterfeiting: Economical Impact on Four Key Sectors of EU Industry*, An econometrics research by CEBR for the Global Anti-Counterfeiting Group, London: CEBR, 2000.

Centre for Economics and Business Research, *Counting Counterfeits: Defining a Method to Collect, Analyse and Compare Data on Counterfeiting and Piracy in the Single Market*, Final report for the European Commission Directorate-General Single Market, 15 July 2002.

Chiou, J.S., C. Huang, and H. Lee, 'The Antecedents of Music Piracy Attitudes and Intentions', *Journal of Business Ethics,* Vol. 57, No. 2, 2005, pp.161–74.

Conner, K. R., and R. P. Rumelt, 'Software Piracy: an Analysis of Protection Strategies', *Management Science*, 1991; pp.125–39.

Cordell, V. V., N. Wongtada and R. L. Kieschnick, 'Counterfeit Purchase Intentions: Role of Lawfulness Attitudes and Product Traits as Determinants', *Journal of Business Research,* Vol. 35, No. 1, 1996, pp.41–53.

Danaher, B., S. Dhanasobhon, M. Smith, R. Telang, 'Converting Pirates without Cannibalizing Purchasers: The Impact of Digital Distribution on Physical Sales and Internet Piracy', *Marketing Science*, Vol. 29, No. 6, 2010, pp. 1138- 51.

Dejean, S., 'What Can We Learn from Empirical Studies About Piracy?', SSRN eLibrary, 2008. As of 9 Mar 2011:
http://ssrn.com/paper=1219442

Depken II, C. A., and L. C. Simmons, 'Social Construct and the Propensity for Software Piracy', *Applied Economics Letters,* Vol. 11, No. 2, 2004, pp.97–100.

De Vany, A.S., and W. D. Walls, 'Estimating the Effects of Movie Piracy on Box-Office Revenue', *Review of Industrial Organization*, Vol. 30, No. 4, 2007, pp.291-301.European Commission, Report on community customs activities on counterfeit and piracy. Results at the European border, 2007. As of 5th September 2011: http://ec.europa.eu/taxation _customs/resources/documents/customs/customs_controls/counterfeit_piracy/statistics2007.p df

European Commission, *Summary report on EU drug precursors seizures*, Directorate General Enterprise and Industry and Directorate General Taxation and Customs Union, 2007. As of 20 July 2012:
http://ec.europa.eu/enterprise/sectors/chemicals/files/new_page/report_2007_en.pdf

European Commission, *A Single Market for Intellectual Property Rights Boosting creativity and innovation to provide economic growth, high quality jobs and first class products and services in Europe*, 2011. Communication, Brussels, 24.5.2011 COM(2011) 287 final. As of 21 July 2012:
http://ec.europa.eu/internal_market/copyright/docs/ipr_strategy/COM_2011_287_en.pdf

EC DG Trade, *Geographical Indications*, 2011a. As of 9 March 2011:
http://ec.europa.eu/trade/creating-opportunities/trade-topics/intellectual-property/geographical-indications/ EC DG Trade, *Intellectual Property*, 2011b. As of 9 March 2011t: http://ec.europa.eu/trade/creating-opportunities/trade-topics/intellectual-property/

EC DG Trade, Intellectual Property: Positive Aspects of IP Rights, 2011c. As of 9 March 2011: http://trade.ec.europa.eu/doclib/docs/2009/january/tradoc_142108.pdf

European Alliance for Access to Safe Medicines, *The Counterfeiting Superhighway*, 2008. As of 26 July 2012: http://www.pharmaceuticalanticounterfeiting.com/uploadedFiles/EventRedesign/UK/2011/February/18399002/Assets/The-Counterfeiting-Superhighway-report(2).pdf

Europe Economics, 'Jobs, tax and contribution to the UK economy: the impacts of reducing illicit peer-to-peer file-sharing', 2008, As of July 2012: http://www.europe-economics.com/publications/fileshare08.pdf

Frontier Economics, *Estimating the Global Economic and Social Impacts of Counterfeiting and Piracy,* Report Commissioned by Business Action to Stop Counterfeiting and Piracy (BASCAP), February 2011.

Forzley, M. 'Counterfeit goods and the public's health and safety', International Intellectual Property Insitutite, 2003.

GAO – see US Government Accountability Office.

Gallup Organization, 'Global Consumer Awareness, Attitudes, and Opinions on Counterfeiting and Piracy', Third Global Congress Combating Counterfeiting and Piracy, 31 January 2007.

Gopal, R.D., and G.L. Sanders, 'International Software Piracy: Analysis of Key Issues and Impacts', *Information Systems Research,* Vol. 9, 1998, pp.380–97.

Grossman, G. and C. Shapiro, 'Foreign Counterfeiting of Status Goods', *Quarterly Journal of Economics*, Vol. 103, No. 1, 1998, pp.79–100.

Hargreaves, I., 'Digital Opportunity', *A Review of Intellectual Property and Growth*, 2011.

Haruvy, E., V. Mahajan and A. Prasad, 'The Effect of Piracy on the Market Penetration of Subscription Software', *The Journal of Business*, Vol. 77, No. 2, 2004, pp.S81-S107.

Harvey, P.J., and W.D. Walls, 'Laboratory Markets in Counterfeit Goods: Hong Kong versus Las Vegas', *Applied Economics Letters,* Vol. 10, No. 14, 2003, pp.883–87.

Hui, KL and I. Png, 'Piracy and the Legitimate Demand for Recorded Music', *Contributions to Economic Analysis & Policy*, Vol. 2, No. 1, Article 11, 2003.

Hunt, Priscillia, B. Kilmer and J. Rubin. *Development of a European Crime Report: Improving safety and justice with existing crime and criminal justice data,* Santa Monica, Calif.: RAND Corporation, 2011.

Hunt, P., L. Rabinovich, and B. Baumberg, *Preliminary Assessment of Economic Impacts of Alcohol Pricing Policy Options in the UK*, Santa Monica, Calif.: RAND Corporation, 2011.

Huygen, A., N. Helberger, J. Poort, P. Rutten & N. Van Eijk, 'Ups and Downs: Economic and Cultural Effects of File Sharing on Music, Film and Games', 2009. As of 9 March 2011: http://ssrn.com/abstract=1350451

IFPI – see International Federation of the Phonographic Industry.

International Federation of the Phonographic Industry, "IFPI online Music Report 2004", 2004. As of July 2011: http://www.ifpi.org/content/library/Digital-Music-Report-2004.pdf

International Federation of the Phonographic Industry, "Digital Music Report 2005", 2005. As of July 2011: http://www.ifpi.org/content/library/Digital-Music-Report-2005.pdf

International Federation of the Phonographic Industry, "Digital Music Report 2006", 2006. As of July 2011: http://www.ifpi.org/content/library/Digital-Music-Report-2006.pdf

International Federation of the Phonographic Industry, "Digital Music Report 2007", 2007. As of July 2011: http://www.ifpi.org/content/library/Digital-Music-Report-2007.pdf

International Federation of the Phonographic Industry, "Digital Music Report 2008: Revolution, innovation, responsibility", 2008. As of July 2011: http://www.ifpi.org/content/library/DMR2008.pdf

International Federation of the Phonographic Industry", Digital Music Report 2009: New business models for a changing environment", 2009. As of July 2011: http://www.ifpi.org/content/library/dmr2009.pdf

International Federation of the Phonographic Industry, *The Impact of Illegal Downloading on Music Purchasing*, 20 November 2009. As of 3 November 2011: http://www.ifpi.org/content/library/The-Impact-of-Illegal-Downloading.pdf

International Trademark Association (1998) The Economic Impact of Trademark Counterfeiting and Infringement. Estimation of the Impact of Trademark Counterfeiting and Infringement on Worldwide Sales of Apparel and Footwear. As of 2 September 2011: http://www.inta.org/Advocacy/Documents/INTACounterfeitingEconomicImpact1998.pdf

IPI – see Institute for Policy Innovation.

Informa Media, 'Informa Media Group US- September 2004' , www.informamedia.com [website unavailable as of July 2012]

Joossens, L., 'Theoretically an option, but an enforcement nightmare'. *Tobacco Control*, Vol. 18, No. 1, 2009.

Klein, B., A. V. Lerner, and K. M. Murphy, 'The Economics of Copyright:"Fair Use" in a Networked World', *The American Economic Review,* Vol. 92, No. 2, 2002, pp.205–08.

KPMG, *Economic Impact Study Analyzing Counterfeit Products in the United Arab Emiriates*, Industry Reports: Brand Owners Protection Group, 2008. As of 5 September 2011 at: http://www.iccwbo.org/uploadedFiles/BASCAP/Pages/Economic%20Impact%20Study%20Analyzing%20counterfeit%20products%20in%20the%20UAE-Final.pdf

Ledbury Research, *Counterfeiting Luxury: Exposing the Myths*, Davenport Lyons, 2007.

Liebowitz, S. J., 'Copying and Indirect Appropriability: Photocopying of Journals', *The Journal of Political Economy,* Vol. 93, No. 5, 1985, pp.945–57.

Liebowitz, S. J., 'Pitfalls in Measuring the Impact of File-sharing on the Sound Recording Market', *CESifo Economic Studies,* Vol. 51, No. 2-3, 2005, pp.435–73.

Liebowitz, S.J., 'Testing File Sharing's Impact on Music Album Sales in Cities', *Management Science,* Vol. 54, No. 4, 2008, pp.852–59.

Liebowitz, S.J., 'The Oberholzer-Gee/Strumpf File-Sharing Instrument Fails the Laugh Test', SSRN eLibrary, 2010. As of 9 March 2011:
http://ssrn.com/paper=1598037

Mahajan, S., 'Concentration ratios for business by industry in 2004', *Economic Trends*, United Kingdom Office for National Statistics, 2006.

Mansfield, E., *Study Guide and Casebook for Managerial Economics*, WW Norton & Company, 1993.

MarkMonitor, *Traffic Report: Online Piracy and Counterfeiting*. Policy Research Report, MarkMonitor, 2001. As of 10 October 2011:
https://www.markmonitor.com/download/report/MarkMonitor_-_Traffic_Report_110111.pdf

Mathers, C. D. and D. Loncar, 'Projections of Global Mortality and Burden of Disease from 2002 to 2030', *PLoS Medicine*, Vol. 3, No. 11, 2006; e442.

Mortimer, H., C. Nosko, A. Sorensen,' Supply responses to digital distribution: Recorded music and live performances', *Information Economics and Policy*, Vol. 24, No.10, 2012, pp. 3-14.

Motion Picture Association of America, *Worldwide study of losses to the film industry and international economies due to piracy,* Industry Report, 2006. As of 14 September 2012:
http://www.iccwbo.org/WorkArea/DownloadAsset.aspx?id=2147491008

Mulligan, M. *Netherlands Recorded Music Revenues Losses Model.* Forrester Research, 2009.

NetResult and Envisional Ltd (2011) 'Update on Digital Piracy of Sporting Events', Submitted in call for Evidence to Independent Review of Intellectual Property and Growth, London: NetResult and Cambridge; Envisional, 2001.

Oberholzer-Gee, F. and K. Strumpf, 'The Effect of File Sharing on Record Sales: An Empirical Analysis', *Journal of Political Economy,* Vol. 115, No. 1, 2007, pp.1–42.

OECD, *The Economic Impact of Counterfeiting and Piracy*, OECD Publishing, 2008. As of 9 March 2011:
http://www.keepeek.com/Digital-Asset-Management/oecd/trade/the-economic-impact-of-counterfeiting-and-piracy_9789264045521-en

OECD (2009) *Magnitude of Counterfeiting and Piracy of Tangible Products: An Update*, OECD Publishing 2009. As of 29 July 2012 :
http://www.oecd.org/sti/industryandglobalisation/44088872.pdf

Olsen, K., *Background Report for the WIPO/OECD Expert Meeting on Measurement and Statistical Issues*, Geneva: OECD Secretariat, 2005.

Pacula, R. L., S. Kadiyala, P. Hunt and A. Malchiodi, *An Alternative Framework for Empirically Measuring the Size of Counterfeit Markets*, Working Paper 18171, National Bureau of Economic Research, June 2012.

Pharmaceutical Security Institute, 'The Spreading Plague of Counterfeiting', unpublished, presentation at the World Bank, 10 March 2005.

Poddar, A., J. Foreman, S. S. Banerjee and P. S. Ellen, 'Exploring the Robin Hood Effect: Moral Profiteering Motives for Purchasing Counterfeit Products', *Journal of Business Research*, 2011.

Qian, Y., *Counterfeiters: Foes or Friends?*, National Bureau of Economic Research, 2011.

Raustiala, K., and C. Sprigman, *The Piracy Paradox Revisited*, Working Paper Series, University of Virginia, 2009, p.140.

Rob, R. and J. Waldfogel, 'Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students'. *Journal of Law & Economics*, Vol. 49, No. 1, 2006, pp.29–62.

Ronkainen, I. A., and J. L. Guerrero-Cusumano, 'Correlates of Intellectual Property Violation', *Multinational Business,* Vol. 9, No. 1, 2001, pp.59–65.

Satchwell, G. *A sick business: Counterfeit medicines and organized crime*. London: Stockholm Network, 2004.

Schneider, M., 'Measuring the Scope and Impact of IPR infringement in the EU;, *Journal of Intellectual Property Law & Practice*, Vol. 7, No. 5, 2012; pp.313–16.

Shapiro, C. and H. Varian, *Information Rules*, Cambridge, MA: Harvard Business School Press, 1999.

Shore, B., A. R. Venkatachalam, E. Solorzano, J. M. Burn, S. Z. Hassan and L. J. Janczewski, 'Softlifting and Piracy: Behavior Across Cultures', *Technology in Society,* Vol. 23, No. 4, 2001, pp.563–81.

Siwek, S.E., 'True Cost of Sound Recording Piracy to the U.S. Economy', Institute for Policy Innovation, 2007. As of July 2012:
http://www.ipi.org/ipi_issues/detail/the-true-cost-of-sound-recording-piracy-to-the-us-economy

Siwek, S. E., *The True Cost of Motion Picture Piracy to the US economy*, Institute for Policy Innovation, Policy Report 186, No. 14, 2006.

Staake, T., F. Thiesse and E. Fleisch, 'The Emergence of Counterfeit Trade: a Literature Review', *European Journal of Marketing*, Vol. 43, No. 3/4, 2009, pp.320–49.

Stryszowski, P., and D. Scorpecci, *Piracy of Digital Content*, OECD, 2009.

Swiss Federal Institute of Intellectual Property, *Counterfeiting and Piracy: An Appraisal of the Current Situation in Switzerland*, 2004.

TERA Consulting, 'Building a Digital Economy: the Importance of Saving Jobs in the EU's Creative Industries', 2010.

Thompson, W.C. 'Bootleg Billions – The Impact of the Counterfeit Goods Trade in New York City', Office of the Comptroller, 2004. As of 20 Jul 2012:
http://www.comptroller.nyc.gov/bureaus/bud/04reports/Bootleg-Billions.pdf

Treverton, Gregory. F., Carl Matthies, Karla J. Cunningham, Jeremiah Goulka, Greg Ridgeway and Anny Wong, *Film Piracy, Organized Crime, and Terrorism*, Santa Monica, Calif.: RAND Corporation, MG-742, 2009.

TRIPS 1994, *Agreement on Trade-Related Aspects of Intellectual Property Rights*, 1994. As of 9 March 2011: http://www.wto.org/english/tratop_e/trips_e/t_agm0_e.htm

Tschmuck, P., *Creativity and Innovation in the Music Industry*, 2nd edition, Heidelberg: Springer, 2012.

US Chamber of Commerce, 'Counterfeiting and Piracy in Brazil: The Economic Impact', 2006. As of 20 July 1012:
http://www.iccwbo.org/WorkArea/DownloadAsset.aspx?id=2147491131

US Government Accountability Office, *Intellectual Property: Observations on Efforts to Quantify the Economic Effects of Counterfeit and Pirated Goods*, GAO Report 10-423, 2010. As of 25 January 2011:
http://www.gao.gov/new.items/d10423.pdf

US International Trade Commission, *China: Intellectual Property Infringement, Indigenous Innovation Policies, and Framework for Measuring the Effects on the US Economy*, Investigation No. 332-514, USITC Publication 4199 (amended), 2010.

Van Barneveld, R., 'Awareness, Perception and Impacts of Internal Market Policies: Results of Eurobarometer Studies in the Area of Intellectual Property Rights', *Presentation at the First Joint Plenary Meeting of the EU Observatory on Counterfeiting and Piracy*, 10 June 2010.

Vida, I., 'Determinants of Consumer Willingness to Purchase Non-Deceptive Counterfeit Products', *Managing Global Transitions*, Vol. 5, No. 3, 2007, pp.253–70.

Waldfogel, J., 'Music File Sharing and Sales Displacement in the iTunes Era', *Information Economics & Policy,* Vol. 22, No. 4, 2010, pp.306–14.

Wall, D.S., and J. Large, 'Jailhouse Frocks: Locating the Public Interest in Policing Counterfeit Luxury Fashion Goods', *British Journal of Criminology,* Vol. 50, No. 6, 2010; pp.1094–16.

Watters, P.A., R. Layton, R. Dazeley, 'How Much Material on BitTorrent is Infringing Content? A Case Study', *Information Security Technical Report*, Volume 16, No. 2, 2011, pp.79-87.Yang, D., 'Culture Matters to Multinationals' Intellectual Property Businesses', *Journal of World Business,* Vol. 40, No. 3, 2005, pp.281–301.

Yang, D., 'The Impact of Business Environments on Software Piracy', *Technology in Society,* Vol. 29, No. 1, 2007, pp.121–41.

Zentner, A., 'Measuring the Effect of File Sharing on Music Purchases', *Journal of Law and Economics,* Vol. 49, No. 1, 2006, pp.63–90.

APPENDICES

Appendix A. A detailed review of approaches to measure IPR infringements

In this section, we focus more closely on the detail regarding the methodologies, assumptions and limitations of existing research reviewed thus far that attempts to estimate the scope, size and effects of counterfeiting and UUPC.

Attempts to measure counterfeiting and UUPC

OECD (2008)

The most comprehensive and widely cited study on the scale, scope and impacts of counterfeiting and UUPC to date is the OECD's *The Economic Impact of Counterfeiting and Piracy*, which was released in 2008. Data for this study were drawn from the UN Comtrade database (for import statistics) and from seizure statistics, assumed to be indicative of the relative intensities of counterfeiting and piracy (C&P) by type of goods and the relative importance of countries as C&P exporters. The main idea underlying the OECD's methodological approach lies in estimating the propensity with which infringing good *g* is imported from country *c* and applying it to international trade statistics to measure the size of counterfeiting and piracy. There are four steps involved in the study's approach:

1) identification of sensitive goods based on the 96 product chapters of the Harmonized System – this will tend to overstate scope as not all products in a chapter are subject to counterfeiting and piracy

3) identification of source economies from customs' surveys and the European Commission DG TAXUD database,[36] where a source economy can be a producer economy or a port of transit economy

4) estimation of counterfeiting and piracy propensities – as customs data is incomplete, propensities for counterfeiting and piracy goods are estimated separately from propensities for counterfeiting and piracy in countries to boost sample size, and then they are combined

---

[36] These are described as 'statistics of customs interceptions' at the external borders of the EU 'of articles suspected of infringing intellectual property rights' on the DG TAXUD website. As of 1 Mary 2012: http://ec.europa.eu/taxation_customs/customs/customs_controls/counterfeit_piracy/statistics/ .

5) application of the good-by-country index to the statistics on imports of good g from country c into country i to get an estimate of the value of counterfeiting and piracy in country i for good g from country c.

The idea is to weigh seizure percentages by shares of total trade in order to get a measure of the intensity of counterfeiting and piracy:

- Counterfeiting factor for goods: seizure percentage of good g is divided by the import percentage of good g.

- Counterfeiting factor for source economies: seizure percentage from country c is divided by the import percentage from country c.

   It is assumed that there is a positive correlation between these factors and the actual number of counterfeit products imported. However, it is also assumed that this correlation might be violated: profiling or ease of detection can lead to higher seizures for products or from countries that represent a lower share of counterfeiting, and lower seizures for those that in fact account for a larger share of the illicit activity. Therefore lower factors may under-estimate actual counterfeiting and piracy. A logarithmic transformation is thus applied, which gives higher relative weight to lower factors. Correction factors are applied to take into account potential product categories and/or source economies that were not identified by the surveys.

- The good-by-country index is computed as the product of the two indexes times a factor denoting the maximum average counterfeit export rate of any type of infringing good from any trading partner.

The OECD (2008) study has the merit of developing an approach that employs data on international trade and statistical techniques to try to link seizure data to observable dimensions. This is a necessary step towards making use of seizure data to measure the magnitude of counterfeiting: given that the underlying distribution of the flows of counterfeits is not known, while it is known that seizures are not a random sample, one has to come up with some clever way of mirroring the customs data into some fully observable dimension. Furthermore, the OECD approach has the merit of being potentially able to provide estimates across all traded goods under the same methodological umbrella. Alongside this significant effort, the OECD (2008) study is also upfront about assumptions underpinning the estimates, and the limitations thereof. The OECD itself admits that 'The overall degree to which products are being counterfeited and pirated is unknown and there do not appear to be any methodologies that could be employed to develop an acceptable overall estimate' (p.71). Table A–1 summarises the main assumptions underlying the OECD estimation procedure.

Table A–1:   Assumptions in The Economic Impact of Counterfeiting and Piracy (OECD 2008)

Type	Assumption	Explicit or implicit?
Distribution of counterfeit and pirated products across countries	It is constant across countries. While the final index is at the good-by-country level, the good index does not vary at the country level, while it is plausible that the incidence of different types of counterfeit goods is different across countries.	Explicit
Alpha parameter[37]	It is equal to10%: this by construction implies that at most 10% of the imports of good g from country c can be C&P.	Explicit, but not well motivated as to why this is the right value
Value	Landed customs value = value of merchandise assigned by customs officials. In most instances, this is the same as the transaction value appearing on accompanying invoices. It includes the insurance and freight charges incurred for transportation.	Explicit
Epsilon parameters[38]	They are equal to 0.05: this means, because the index is in logs, that the propensity of a given good and the propensity of a given economy are adjusted upwards by 5% in order to account for C&P that could not be identified by the surveys.	Explicit, but not well motivated. It seems likely that this could be good- and economy-specific instead of being constant.

Table A–2 summarises the limitations of the study.

Table A–2:   Limitations of The Economic Impact of Counterfeiting and Piracy (OECD 2008)

Limitation	Why	Implication
The adjustment factors (alpha and epsilons) applied to the indices are constant.	The distortions they adjust for are assumed to be constant over time and across regions.	There is no attempt to consider differential market susceptibility to C&P, which is highly likely, so policy implications will be severely limited and estimates of economic impact can be biased (direction of bias is unclear).
Incomplete customs data: some countries do not respond at all, some do not indicate products, some do not indicate sources.	It is not possible to estimate a complete matrix of good by source country factors, so in order to boost sample size good and source country propensities are estimated separately.	The authors have to assume that the distribution of counterfeit and pirated good types is constant across countries.
The distribution of counterfeit and pirated products is assumed to be constant across countries.	This is a very strong assumption, which is equivalent to assuming that the underlying market characteristics are constant across countries (obvious evidence contradicting this assumption is differences in prices and income across countries).	There will be bias in the results if the assumption does not hold, upwards if the distribution overstates the importance of higher-value goods, downwards otherwise.
The methodology covers imported goods only.	This methodology does not cover counterfeit and pirated goods that are produced and consumed domestically.	The methodology ignores the total economic impact of the problem.
Coverage is incomplete.	This methodology does not cover digitally distributed products.	The methodology ignores the total economic impact of the problem.

---

[37] This is the 'maximum average counterfeit export rate of any type of infringing goods originating from any trading partner' (p.131).

[38] These are correction factors applied during the estimation process to account for source economies and/or product categories that for any reason might not have been captured by customs data.

The USITC (2010) report critiques the OECD approach, characterising the use of seizure data as problematic – one of the problems they notice is that customs act only on products that are registered with them, and the percentage of registered trademarks and copyrights that holders actually notify to customs is generally low, so that a lot of products might be missing for non-random reasons from the picture that seizure data provide; in fact the European Commission report (2007) mentions that 80 percent of their seizures are done based on some kind of signalling by the right holders. Another legitimate point that can be made about using seizure data at the border is that such data concern only counterfeits that are in transit. If the products' final destination is outside the country in which they are seized, then inferences based on seizure data will account for each transit across borders and thus over-estimate. Bate (2008) reports that it is a common practice to ship counterfeit medicines to intermediaries in the US and UK before they reach their final destination.

Frontier Economics (2011)
In a recent report commissioned by the International Chamber of Commerce-Business Action to Stop Counterfeiting and Piracy (ICC-BASCAP), Frontier Economics extended the OECD (2008) methodology by adding estimates of domestic counterfeiting and piracy and of digital piracy. Furthermore, in building on the OECD work, Frontier Economics (2011) places greater weight on the importance of the aspects that the OECD had not quantified, namely domestically-consumed counterfeits and pirated goods.

The report does not develop a conceptual framework as such – rather, it implicitly adopts the OECD framework, given that its main exercise is an update and extension of the OECD estimates. As in the OECD (2008) study, the Frontier (2011) estimation exercises are from a supply-side perspective in that they build off of statistics for supplying C&P goods to the market (e.g. seizure statistics, trade statistics).

There are three main stages of the estimation by Frontier Economics:

1) Assume that 25%–75% of the increase in seizures relative to imports is due to counterfeiting and piracy. This is equivalent to saying that part of the relative increase in seizures can be attributed to increased detection ability by customs. The additional value of seizures due to counterfeiting and piracy is obtained by multiplying the increase in seizures relative to imports by 25%–75% and added to the $250 billion OECD estimate to get an updated estimate of internationally traded counterfeiting and piracy.

6) Estimate the value of domestic counterfeiting and piracy by applying the counterfeiting propensities for each product category identified by the OECD to national GDP statistics through the link between product categories and GDP components.

7) Estimate the value of digital piracy based on industry and academic studies.

This study is appreciable in its effort to fill in the gaps of the OECD (2008) study, i.e. quantifying domestic counterfeiting and piracy. However, in order to expand the scope of the OECD (2008) study both in time and in content it has to make further assumptions and it incurs a number of new limitations. Table A–3 presents the nature of the new assumptions adopted in Frontier Economics (2011) beyond those incorporated in the OECD study. Table A–4 lists the limitations of the approach and their implications.

Table A–3:   Additional assumptions in *Estimating the Global Economic and Social Impacts of Counterfeiting and Piracy* (Frontier Economics 2011) compared to the OECD (2008) study

Type	Assumption	Explicit or implicit?
Dynamic relationship between seizures and counterfeiting	25–75% of the increase in seizures relative to imports is due to counterfeiting and piracy.	Explicit
Relationship between C&P and exports.	It is constant across countries.	Explicit, and partially relaxed drawing on other literature
Price	Price is average domestic market price of a given product category because domestic GDP is used to compute domestic C&P.	Implicit

Table A–4:   Additional limitations of *Estimating the Global Economic and Social Impacts of Counterfeiting and Piracy,* (Frontier Economics, 2011) compared to the OECD (2008) study

Limitation	Why	Implication
Extrapolation	Applying just a linear transformation to the seizure data for 2005–2008 is restricting the functional form of a relationship which might well be non-linear. (Very small increases in seizures could be due more to higher incidence of C&P than to higher enforcement, very high increases could be due more to higher enforcement than to higher incidence).	The updated estimates are biased if the assumption does not hold. Bias can go in either direction.
Estimation of domestic counterfeiting	The propensities that are used for the estimation were derived by the OECD using data on external trade. The same caution that the OECD used in not applying such propensities to domestic product applies here.	If the composition of products in the domestic production and consumption of counterfeits is different from the composition of products in the international trade of counterfeits then the estimates are biased.
Calculation	The straight summation of two estimates obtained with very different methodologies has to be done with great caution.	The Frontier Economics update part is likely to be less precise than the OECD baseline estimate because it relies on strong assumptions and less fine-grained data.
Transparency	There is no mention of the details of the regression analyses used to extrapolate data for a larger number of countries.	The method cannot be replicated.
GDP breakdown	It is assumed that manufacturing at large is a target of C&P but that might not be true for certain product categories.	Over-estimate.

CEBR (2002)

In a study for the European Commission Directorate General Internal Market and Services (DG MARKT), the Centre for Economic and Business Research (CEBR) (2002) developed a general methodology for data collection on counterfeiting (and piracy) that could improve efforts to construct demand-side estimates of the size of the problem. The CEBR report suggests the integration of various demand-side data sources to generate C&P estimates and to rely on seizure data only if detection rates are known and known to be high. While there is no stated improvement of modelling beyond the data, the methodological contribution is in the triangulation of information obtained from different data sources.

The CEBR study provides an overview of different possible data collection methods. In addition, the trade-offs between breadth of applicability and costs are presented. A decision tree is developed based on the characteristics of the product categories, which leads to the choice of a data-collection methodology. Specifically, authors identify relevant nodes as the following:

- level and confidence of detection rates;

- prevalence of extra-EU trade vs. domestic;

- distribution of producers and suppliers;

- awareness of suppliers and consumers.

CEBR provides recommendations for the ways in which to collect data, by product type (see Table A–5). One of the key criteria for the choice of the recommended methodology is based on whether the consumer can distinguish a fake from the original, which has implications for the assumptions on the substitution rate of the genuine product for the counterfeit.

Table A–5: Recommendations for data collection to estimate the size of counterfeiting markets (CEBR, 2002)

Data source recommended	Product type
Consumer surveys	Books, electronic games, movies, music, sunglasses and watches
Consumer surveys and mystery shopping	Clothes, perfumes, leather goods and spectacles
Supplier surveys	Motor vehicle, aircraft and other industrial spare parts
Supplier surveys and mystery shopping	Alcohol, food, pharmaceuticals and plants

The CEBR study recommends also conducting consumer surveys to elicit self-reports of C&P purchases when relying on data collected through mystery shopping in order to be able to weight for the representativeness of different outlets in which mystery shopping occurs. Furthermore, the report argues that in the case of production over-runs,[39] neither mystery shopping nor surveys can effectively identify counterfeiting.

Lastly, a methodology is developed for filling in missing data. Essential characteristics of product categories and country pairs are identified and ranked, including production costs, barriers to entry, enforcement and penalties. A propensity to counterfeit (or pirate) is calculated as a weighted average of the rankings, then estimates for a pair x,y are weighted by the ratio of the propensity of another pair c,d over the propensity of pair x,y to get the estimate of pair c,d. There is a large literature on methods to address missing data; for a bibliographic list of articles and methods, see Hunt, Kilmer and Rubin (2011).

The study lacks transparency in addressing one of the key methodological questions: the reliability of seizure data through the assessment of detection rates. Such lack of transparency weakens the case for a demand-side approach. In fact, the key node in the decision tree to exclude the use of seizure data is the question of whether the detection rates are known with confidence and known to be high. However, the study does not mention which information was used to qualify the confidence level attached to detection rates, nor what the cut-off should be in order to consider them 'high'.

Once the case for the use of consumer and supplier surveys is made, one merit of the study is not to neglect the difficulties of such approaches. The limitations of the quantitative elicitation of consumption patterns through self-reports are discussed in detail. Obviously there is a tendency to under-report any behaviour that is illegal or considered to carry social stigma. Computer-assisted self-interviews are recommended to improve response rates and robustness over other

---

[39] Cases in which the over-run production of legitimate production sites is distributed to wholesalers and/or retailers without consent of the IPR holder.

modes. The need to ensure demonstrable anonymity is also stressed. The study makes a significant methodological contribution in identifying goods for which the integrative use of mystery shopping information can capture an important aspect of deceptive counterfeiting that cannot be assessed through surveys.

KPMG (2008)

KPMG released a study in 2008 in which the size of counterfeiting in four sectors of the United Arab Emirates economy is estimated. The client for this study was the Business Owners Protection Group (BPG), which represents some of the largest trademark owning companies in the country. The data collection was conducted by BPG; no further details are provided, but the outcome of the process is market shares of counterfeiting in the different sectors. These were applied to the values of the respective markets to derive estimates of the size of counterfeiting. This calculation assumes a substitution rate of 1, ie all those who buy counterfeits would buy original goods in absence of counterfeiting. This is a questionable assumption and together with the lack of transparency about how the single industries derive a point estimate of the incidence of counterfeiting in their respective markets, calls the overall architecture of the study into question. Any further assessment is impossible without knowing more about the data-collection process. This appears as a significant weakness for a study in a sector in which the major problem is the lack of reliable data.

International Trademark Association (1998)

A number of studies try to estimate the value of sales lost by IPR holders as a proxy for the value of counterfeiting. This appears to be a more immediately applicable measure from the perspective of legitimate businesses in the policy process, but it is also not immune to limitations. One of the earliest studies using this approach was conducted by the International Trademark Association (INTA) in 1998, focusing on footwear and apparel. The authors postulate a relationship such that total revenues (TR) in a given market are a function of trademark protection (P) and a set of controls Z (GDP per capita, population). They pool data from different countries over multiple years to boost sample size. Data are collected on perceptions of trademark protection functioning (through a survey of 230 INTA members in 40 countries) and on sales revenues (through a survey of 46 INTA members asked to provide their sales data for as many as possible of the 40 countries from 1991 to 1995). INTA develops a sort of reduced-form model that relates revenues directly to trademark protection. In contrast, a structural model would require specifying the relationship between trademark protection and demand for counterfeit products. In the INTA 1998 model, the elasticity of sales to trademark protection is estimated to be -0.25 in both sectors over 1991−1995.[40] The percent loss of sales (X) is thus calculated as:

$X = (P^{-0.25} - 1)$

By construction, the percentage loss of revenue can be at most -33 percent because P is only allowed to vary from 1 to 5. The estimated percentage losses in sales are then applied to the sales data.

Aside from the small sample size and relatively short time-range which limit the significance of the results, the study seems to lack transparency in at least two key areas. First of all it is not clear whether the sample of companies surveyed is random or representative of the products and

---

[40] The negative sign is due to the fact that the trademark protection variable is highest for lowest protection.

brands most likely to be counterfeited. There may be selection biases in this sample, which would make the inferred estimates not representative of the population. Secondly, the choice of the variables of the model seems rather arbitrary: it is not clear, for instance, what theoretical considerations suggest that the impact of trademark protection on sales is a function of population (the study reports that the coefficient on the trademark protection variable is significant only when interacted with population). Nonetheless, the mathematical structure offered to describe the relationship between revenues and trademark protection represents a useful contribution for moving the field forward.

Swiss Federal Institute of Intellectual Property (2004)
A study conducted by the Swiss Federal Institute of Intellectual Property (SFIIP) (2004) asked companies in a survey to estimate the losses incurred because of counterfeiting. Theoretically, legitimate suppliers should be able to provide a reasonably good gauge of lost expected revenue if they have models that help them predict sales in a given market based on (and controlling for) typical market factors and fluctuations. If the models are reasonably good at forecasting, deviations from these predictions could be presumed to be due to counterfeiting. Similarly, producers might be aware of the level of counterfeiting because of legal suits they are engaged in related to patent or copyright infringement.

However, a key disadvantage of eliciting an estimate of counterfeiting from suppliers, particularly in studies such as this one, is that frequently no indication is given regarding how the respondent has come up with the loss estimate they have provided. In the SFIIP (2004) report occasional references are made to some typical reference values (such as the amounts awarded by judicial authorities as compensation acting as a lower bound for the estimate). Moreover, if there is variation in the benchmark used by suppliers to report counterfeiting, and differential bias associated with different approaches, then the use of multiple methods for benchmarking could be problematic. In the specific case of the SFIIP study, there was a very high non-response rate to this specific question, which together with the small sample size (72 companies) and the high variance in the distribution of responses (which range from tens of thousands of Swiss francs to annual losses of 800 million Swiss francs in the watch industry) makes the presumed accuracy of a population mean highly suspect.

CEBR (2000)
CEBR conducted a study for the Global Anti-Counterfeiting Group in 2000 analysing C&P in four key sectors in the EU: clothing and footwear, pharmaceuticals, toys and sport equipment, and perfumes and toiletries. Revenue losses were estimated starting from likely incidence factors provided by industry associations, applying such factors to actual revenues, and assuming perfect substitution. The authors subsequently used opinion polls to estimate substitution rates to scale down revenue estimates.[41] This is certainly better than assuming a substitution rate of 1, but the fact that substitution rates are set to be constant across countries, while probably driven by data availability, is still not ideal as tastes and practices are likely to vary across borders. In addition, profit losses were estimated applying marginal profitability to revenue losses. Finally, an important limitation of the study is that data on which it is based are reported to be complete

---

[41] The opinion polls were conducted in the UK and Italy. The study does not provide details about the samples. It only reports: 'Opinion polls [were] asking the question below or a variant: "Which, if any, of the following goods would you knowingly purchase as counterfeit, assuming you thought the price and quality of the goods was acceptable?".'

only for the UK, which can introduce significant error if used outside of the UK as there may be substantial cross-country differences.

Allen Consulting (2003)

Using a similar approach to the one by CEBR (2000), the Allen Consulting Group undertook a study for the Australian business associations in the toys, business software and computer and video games sectors. One interesting point made in this study is that determining the size of the legitimate industry is itself problematic because industry classifications are usually not at the level of detail that is needed. An estimate of the net revenue loss to IPR holders as a consequence of counterfeiting is calculated in four steps:

1) An 'incidence' of counterfeiting parameter is taken from other studies (for example, the CEBR (2000) study is used for toys).

2) The incidence is applied to the statistics on the value of sales of the original product to get an estimate for the maximum value of lost sales.

3) This estimate is adjusted with substitution rates obtained from consumer surveys to get an estimate of the value of lost sales.

4) Confidential data on profit margins by businesses are used to generate a net revenue loss.

These steps are not presented explicitly, so estimates could not be replicated even if confidential data on businesses' margins were available. Moreover, the revenue loss estimates are clearly very sensitive to what margins are applied, so that a lot of weight is put on the accuracy of the confidential data provided by the industry.

European Alliance for Access to Safe Medicines (2008)

The European Alliance for Access to Safe Medicines (EAASM) 2008 report contains an estimate of the prevalence of counterfeits in online pharmacies through a mystery shopping exercise. The methodology is composed of three stages:

1) A sampling frame of online pharmacies selling prescription-only medicines was created by querying top search engines with keywords such as 'online pharmacy', 'cheap medicines' and 'medicines online'. The search returned 116 websites. These are typically characterized by no traceable physical presence, lack of licensure or approval from a recognised association or oversight body, and extensive violation of intellectual property.

2) Orders were placed for 36 different medicines based on the list of most sold products in the US.

3) An expert panel of 22 members examined the products based on packaging, patient information leaflet, patient information language and the condition of the blister pack, frequently encountering incorrect or suspect elements. The subsequent laboratory analyses confirmed that 62% of the medicines were substandard or counterfeit.

The mystery shopping methodology appears to be appropriate in this case because consumer ability to distinguish counterfeits from the original products is low. However, the study does not quantify the actual incidence of counterfeit purchase, because it does not take into account that a gradient of awareness levels is likely to exist among consumers, based among other factors on age, familiarity with the Internet and education. So, the probability with which consumers buy from different online retailers might vary based on how trustworthy they judge the retailers to be, in

which case the 62% figure might not be very accurate in representing the actual prevalence of counterfeit drugs. In order to translate this figure into an estimate of counterfeiting, the approach would have to try to somehow capture the heterogeneity in the trade-off between lower prices (and possibly higher utility from not having to disclose in person a particular health condition) and the perception of the likelihood of a counterfeit purchase.

Joossens (2009)

Joossens (2009) published a study on the global illicit cigarette trade. The study notes a distinction between illicit trade (which exploits price differentials due to taxation in order to generate illegal arbitrage opportunities) and counterfeiting of cigarettes. Illicit trade does not violate any IPR. However, although the study notes that part of tobacco smuggling is represented by counterfeit cigarettes – it does not attempt to disentangle the two, leaving the reader with the impression that a practical way of distinguishing them does not exist. Thus estimates of illicit cigarette trade represent an upper bound for counterfeit cigarette trade.

The study updates country-level estimates of the illicit cigarette market around the world, using 2007 data or data as close to 2007 as available. The methodology consists of an aggregation of different country-level estimates from academic articles, official government publications, estimates from market research companies, tobacco trade journal articles, newspaper articles and authors' personal contacts in customs organisations. The resulting bottom-line statistic is an estimate that 11.6% of the global cigarette market is illicit, equivalent to 657 billion cigarettes a year and $40.5 billion in lost revenue. The relationship between income and illicit cigarette trade is also investigated, showing that it goes in the opposite direction to the common wisdom. In fact, higher income countries, where cigarettes are more expensive, have lower levels of cigarette smuggling than lower income countries.

The methodology used to arrive at this estimate benefits from wide geographic coverage (84 countries). However, while representing an appreciable effort to use the available information in the most effective way, it also has limitations because it relies on widely different sources, which means that there is variation in reliability, precision and ultimately comparability of the resulting estimates.

Attempts to measure UUPC exclusively

While many of the studies cited thus far included attempts to measure the size of the market for UUPC in conjunction with counterfeited goods, they employed the same methodology across for both counterfeiting and piracy. Here we review a sample of literature focused exclusively on UUPC, or piracy as it is most commonly termed, enabling us to consider more thoroughly the unique features of the methodology, assumptions and limitations of research examining piracy specifically.

Interestingly, unlike the case for counterfeiting, the methods reviewed focused on piracy tend to be fairly uniform. Many follow the recent GAO (2010) recommendation suggesting that digital products should be estimated separately from other pirated goods because the products are not physical or tangible, they can be reproduced at very low cost, and they have the potential for immediate delivery through the Internet across virtually unlimited geographic markets.

In what follows we summarise key studies estimating the magnitude and impacts of piracy.

TERA Consulting (2010)

A study by TERA Consulting (2010) aims to estimate the direct and indirect impacts of piracy, i.e. revenue losses and job losses as a result of piracy.[42] In order to do this, TERA Consulting calculates the number of infringements in 'core' and 'non-core' industries, which provides an estimate of the size and value of piracy. The 'core' and 'non-core' industries, are defined as those industries that are completely or mainly based on copyright ('core' creative industries) or industries that depend to a lesser extent on copyright-protected materials ('non-core' creative industries). Therefore, the study seeks to capture wider impacts of copyright infringement by considering losses to the copyright owners and to industries which support bringing the product to market.

The industries are categorised in the following manner:

- Core industries
  - press and literature
  - music, video, software
  - databases
- Non-core industries
  - Interdependent industries
    - TV, radio, CD, DVD players
    - computers and equipment
    - musical instruments
    - photographic and cinematographic instruments
    - blank recording material
    - paper
    - other
  - Non-dedicated support industries
    - general wholesale and retailing
    - general transportation
    - telephony and internet

The study also makes a distinction between 'digital' piracy and 'physical' piracy. Digital piracy is described as 'copyright infringement of digital media [and] refers to various forms of online piracy, including file-sharing via peer-to-peer (P2P) networks' (TERA Consulting, 2009, p.4). Physical piracy is described as the 'sale of illegally duplicated and distributed copyrighted physical works (CDs, DVDs)' (ibid., p.18).

---

[42] Whilst this study also covered the impacts of counterfeiting, we have only considered the aspects relevant to piracy in this deliverable. Next iterations of this deliverable will include a detailed review of the methodology and estimates for counterfeiting in this study as well.

In order to calculate size and value of these illicit markets, the study uses country-specific and industry-specific survey results in Europe's five largest markets (the UK, France, Germany, Italy and Spain), which collectively represent approximately 75 percent of European GDP. TERA Consulting first calculates the number of IPR infringements and associated value of infringements due to piracy as the sum of the number of illegally downloaded files, streamed files (films and TV series) and physical counterfeit products. The input data, from numerous sources for each country and industry, were gathered through a variety of methodologies, including surveys, interviews and downloading statistics. These include for example, the *European Video Yearbook 2009*, *Video on Demand in Europe 2008*, *Ipsos Digital & Physical Piracy in GB 2007* and the *Spain Ministry of Culture Report*. There are some citations for which full referencing is not available, and more research is needed to locate the sources and/or verify whether this information can be disclosed.

In order to then translate IPR infringements into revenue and job losses, the study seeks to answer the questions: 'How many of the infringements would have been legal purchases?' and 'What would have been the price paid for the legal product?'.

The potential number of lost legal purchases is calculated by using a rate of substitution of pirated material for authorised material. The substitution rate in this study is determined in two ways due to data limitations:

- the change in the probability of purchasing the legal product;
- the change in the number of purchases (due to the consumption of counterfeits or a service providing pirated products).

Prices paid for legal products are acquired from surveys and reports identified by TERA Consulting (2010) and derived through a number of methodologies. For an item of TV broadcasting content, for example, the price is based on the average time spent watching TV per day and millions of Euros spent on TV ads.

Each of the substitution rates and prices of pirated products used to calculate impacts is assumed to depend on the exact product or level in the supply chain. Table A-6 shows the assumptions for each product at various stages of the supply-chain identified by TERA Consulting (2010). As an example, for a film that first enters the cinema and flows through a number of channels as a digital product, the percentage substitution rates are: 5, 10, 0, 0, 10 and lastly 10. There are specific prices for each of these stages. Another example is music. It is assumed that of the number of infringements made online (i.e. 'digitally') 10 percent would be purchased legally and of those 'physical' infringements, 45 percent would be purchased legally.[43] In other words, people purchasing pirated CDs are more likely to purchase music legally than those who acquired pirated music in its online forms. The prices then applied to the digital infringements are 90 percent of the online price of music plus 10 percent of the average price of a CD. For physical piracy, 100 percent of the average CD price is used.

---

[43] Largely, these numbers of infringements appear to be estimates from the International Federation of the Phonographic Industry (IFPI) based on Jupiter Research. There is one full reference provided for IFPI and further research will be carried out to identify these values and methodologies.

Table A–6:   Substitution, price and consumer deception assumptions in the TERA Consulting (2010) study

Type	Assumption	Explicit or implicit?
Substitution rate	The study assumes fixed rates distinguishing a number different piracy product categories for both digital and physical piracy: • music • films: in cinema, on TV • TV series: video-on-demand, pay-per-view	Explicit
Price	The study distinguishes different price of between types pirated products calculated from the price of the legitimate products: • for music (online): 90% of the digital music price + 10% of the CD price • for music (physical): 100% of the CD price • for film & TV (online and physical): specific to the stage in distribution (digital & physical the same)	Explicit
Consumer deception	High deception. The study assumes that consumers have little awareness of whether they are purchasing a legal or a pirated product.	Implicit

Lastly, the findings on revenues and job losses for the five countries are then extrapolated to the EU level. This is performed in two steps. First, it is assumed that the revenue losses due to piracy in other EU countries are proportional to their share of total GDP in the EU. This proportion is directly available for core industries, but not for the non-core industries. Therefore, in a second step, to calculate for the non-core industries, a weight is applied to the proportion. The weight is calculated as the ratio of trade value of core industries to non-core industries at the European level (not the country level because of data gaps).

This study takes into account an important aspect of the market for pirated (and other) goods: the supply-chain. In particular, the distribution of a pirated product can occur when a legal consumer becomes an unauthorised distributor without moving the product physically through the supply-chain. As such, digital and physical pirated products are likely to have different rates of substitution at different stages. Equally, prices may differ for reasons described by the GAO (2010), namely the product type, packaging and location of distribution.

This study also takes into account a number of surveys at country- and industry-specific levels. An attractive feature of this approach is that the substitution and prices observed in the market are likely to depend on consumer deception and product quality (implicit assumptions from substitution and price) that differ by industry and country.

In addition to these strengths, there are a few key weaknesses. First, the extrapolation from five large EU member states to the other member states using the proportion of their share of total GDP in the EU may be seen as a limitation since it is not clear that piracy levels are proportional to legitimate sales volumes. Second, a large number of sources is used for prices and substitution rates, which makes it infeasible to review the data quality of all those sources. As such, it becomes a difficult study to assess and replicate.

IFPI (2004-2009)

The International Federation of the Phonographic Industry (IFPI) has published a series of annual reports since 2003 discussing industry events over the preceding year, including statistics on the state of the industry. While no information regarding methodology is provided (as these reports are intended for a less technical, and large, audience) it appears, that perhaps a consumer survey is conducted to identify the number of IPR infringements in music (IFPI, 2004, 2005, 2006, 2007, 2008, 2009). It is, for example, unclear whether certain information is derived from surveys conducted by IFPI, surveys commissioned by IFPI and/or other literature cited by IFPI.

References are not provided. Furthermore, for most statistics provided, it is unclear who is surveyed in the population or the sampling frame.

Mulligan (2009)

A study carried out for Forrester Research (Mulligan, 2009) assesses the scale of recorded music piracy using consumer survey data for P2P and non-network sharing, scaling up the numbers in order to account for under-reporting (no details are provided on how this is done). In order to estimate forgone sales, the study makes some assumptions about the substitution rate and average price, although again no details are provided. In order to quantify lost revenues the study assumes an exogenous composite annual growth rate of 0.8%, so the 'lost growth' is added to the lost revenue. An important contribution of this study is that it considers lost growth, which has not been explicitly considered in any of the other studies reviewed here. However, lack of transparency regarding the estimates makes the results difficult to use.

Institute for Policy Innovation (IPI) (2007)

The IPI 2007 report is also on the recorded music industry. The way the authors quantify the size of piracy, draws on an IFPI report (2006) called *2006 Global Recording Industry in Numbers* which estimates the physical piracy rate by country. IPI uses this information to compute pirated units as the product of the piracy rate and legitimate sales, and lost sales as the product of pirated units and the retail price.

IPI (2007) down-weights this measure by several factors:

- First, the substitution rate is computed as the ratio between the number of original units that would be sold in absence of piracy and the estimated number of pirated units sold. The counterfactual (i.e. the number of original units that would be sold in absence of piracy) is estimated as the ratio between the sales of pirated units and the price of the original units. The resulting substitution rate is 65.7%, in line with the range the authors report to have found in surveys (between 40 percent and 70 percent).

- Second, lost revenues are down-weighted by the share of works whose IPR holder is US-based.

- Finally, lost revenues to industry are separated from lost revenues to retail by weighting the estimate by the relative return (the ratio of trade price to retail price to get the return to the industry; the ratio of the difference between retail price and trade price to retail price to get the return to the retail sector).

For digital piracy, the authors take the number of illegal downloads from the 2006 IFPI report, down-weight it by the share of US repertoire and by a substitution rate of 20% based on a literature review.

This report is very rigorous in documenting its sources and describing its assumptions. It goes beyond the average level of detail found in other studies, taking into account the specificity of the US market and explicitly distinguishing between losses to different market players (sound recording industries and retailers). Its main merit lies in the ability to translate the 'crude' IFPI figures into estimates of lost revenues. However, this requires an additional, complex set of assumptions on prices that entails several moving parts (prices of different products, weight of different products, relative weight of a song within a CD). Sensitivity analysis would be useful to quantify the relative importance of these assumptions.

BSA-IDC (2010)

The Business Software Alliance published a study in 2010, conducted by International Data Corporation (IDC), estimating the global commercial value of pirated software. In deriving the size and value of unlicensed software in a country in a given year, the study uses a number of steps:

1) determining how much PC software was deployed in a specific country in 2009;

2) determining how much PC software was paid for/legally acquired in that country in 2009;

3) subtracting one from the other to get the amount of unlicensed software;

4) dividing by the average system price to get the commercial value of pirated software.

Steps 1-3 are performed to calculate the size of software piracy; step 4 calculates the value. In step 1, BSA-IDC (2010) uses survey information to calculate the total software units installed. The calculation is the number of PCs with the software multiplied by the units per PC.

For countries that are not surveyed, the study uses a methodology that relies on a correlation between factors. This methodology is not provided, but IDC states that it tracks trends in the computer industry, and performs annual surveys of the markets in those countries, targeting vendors, end-users, etc. The authors also state that IDC has local analysts to provide further details on the market. For the countries with severe data limitations the study assumes that conditions are equal to similar countries. However, details on this extrapolation methodology are limited.

In step 2, the study uses market data and annual research to estimate the total number of legitimate software units. This is calculated as the total value of the software market divided by the average system price. Step 3 subtracts the amount of software on PCs (step 1) from the amount of software sold for PCs (step 2). If this is equal to zero, for example, then there is no piracy – all software installed on PCs was legally acquired. If the amount is positive, there is piracy.

For step 4, BSA-IDC uses the following elements that were required to calculate steps 1 and 2:

- the number of total units of software installed;

- the number of legitimate units of software installed;

- the number of unlicensed units of software installed;

- the average system price for legitimate software.

Specifically, the value of the unlicensed software is assumed to be equal to the number of unlicensed software units multiplied by the average system price.

A potential weakness with this approach is in the specification of the average system price. The authors state 'in practice, because of the many methods of deploying software, the average system price is lower than retail prices one would find in stores' (BSA-ID, 2010, p.15). This would require further investigation since the highest piracy rates tended to be in poor countries – Georgia, Zimbabwe, Bangladesh, Moldova, Armenia and Yemen – where it is very likely that even legitimate goods would not be purchased (and therefore priced in the legitimate market) at system price.

Moreover, the study provided limited transparency as to how the piracy rate is determined. IDC indicates that the PC software piracy rate for a given country reflects a complex set of inputs to the simple equation that produces the rate. These include:

- PC shipment growth

- activity in the installed base of older machines

- consumer versus business ownership

- distribution channels — especially growth or decline of non-branded vendors

- legalisation and special pricing programs of vendors

- availability of legitimate software

- availability of pirated software

- broadband access

- desktop-to-laptop mix

- economic conditions, taxes and exchange rates that affect software prices or the discretionary income of buyers.

The relations between these variables, however, are not specified. The equations underpinning these relations are important because they are likely to be used for the high piracy countries. There is no consensus in the literature about how some of the elements used in this extrapolation (e.g. broadband access) influences piracy rates.

Other studies on online piracy
Another group of studies on online piracy uses monitoring of online traffic as an input to estimate the volume of online piracy. Although they constitute a relevant source of information, we concentrate less on them as they are based on methodologies that are not directly replicable, either because they focus on unique events or because they rely on proprietor technology.

For example, NetResult and Envisional Ltd (2011) published a report on the online piracy of live sport events. Its methodology is based on case studies that monitor the supply of illicit online broadcasting of particular sport events or leagues in football, cricket, tennis, Formula One, golf and rugby. It is thus based on primary data collection and is not replicable, as no detail is provided on how the sources of streaming were identified (e.g. search criteria and search engines) and counted. One important market characteristic that emerges from this study is that the illegal streams have come to perform really close to the original broadcasts in terms of quality. MarkMonitor (2011) published a report on online intellectual property theft that uses data on internet traffic. This is applicable both to digital piracy and to online sales of counterfeit goods. The report makes use of a proprietor technology that allows the selection of a list of brands, scanning of websites, their classification and a count of visits. Attributor (2010) estimates the demand for pirated e-books through Google AdWords, which counts the number of clicks generated by keywords. They also have a proprietor technology to capture the number of successful downloads per title, then link it to price data from Amazon.

The literature on estimates of UUPC is only indicative; many other empirical papers exist which present some measure of UUPC or which attempt to show the association between UUPC and other phenomena, such as legitimate music sales and concert ticket sales, legitimate DVD purchases, box office revenues, and so forth. Examples of these include Danaher *et al.* (2010);

Bhattacharjee *et al.* (2006); Mortimer *et al.* (2012); Watters *et al.* (2011); Waldfogel (2010) and De Vany and Walls (2007) among others.

Appendix B. Non-technical description of a new method to measure IPR infringements

The information included in this appendix was sent out to members of the European Observatory on Counterfeiting and Piracy with the aim of recruiting potential firms to participate in the piloting the methodology proposed in this report. We prepared a brief non-technical description of the approach and a series of common questions and answers. It should be noted that the description in this appendix is not strictly consistent with the methodological explanation in Chapter 5. This is because its features have slowly evolved over the course of the study.

Introduction

Researchers with RAND have been working by contract for the European Commission, with the goal of developing better methods for estimating the impact of counterfeit and unauthorised use of protected content (UUPC) in the European Union. A key aspect of this work is a methodological framework for improving estimates of counterfeiting and UUPC for products in key sectors. As many are aware, UUPC and counterfeiting are serious concerns because of the economic, social and health impacts consumption of these unauthorised products can have.

The RAND team has developed a promising approach, based on the findings of an extensive review of existing methodologies and relevant literature. The model that is developed builds on the lessons from more than 250 sources. The approach has been reviewed by professors of economics and criminology across Europe and the US, as well as by analysts from private firms that are responsible for estimating counterfeiting and UUPC within these firms. While legitimate concerns have been raised regarding the level of detail to include in the RAND model within specific industries, the feedback has been generally positive and optimistic about the approach. Those questions/concerns that remain are issues that can only be understood, evaluated, and addressed through the application of the model on real firm data from multiple industries. Thus we currently seek to apply our methodology to relevant firm and/or industry data from a variety of sectors to evaluate, validate and improve the model.

A brief explanation of the approach

While it is well understood that every firm and product is different and has specific nuances that make estimation of counterfeit and UUPC unique, some of these issues can be accounted for statistically when data from multiple firms are merged together into a single empirical model. Some of the issues can only be dealt with by modifying the model or developing tiered models. The pilots help us to better understand these aspects, and how to take account of these market specificities. Our most general model makes use of well-understood firm behaviour derived from economic theory to approximate the unexplained annual change in unfulfilled demand for legitimate product. The unexplained unfulfilled demand refers to that amount of product that was not sold and could not be explained retrospectively by firm specific supply chain factors, industry factors or other market shocks that became realized during the evaluation period. This unexplained unfulfilled demand, our model presumes, is due at least in part to counterfeiting and UUPC. The model does not presume all of it is counterfeiting and UUPC nor does it try to calculate all the reasons legitimate demand was unfulfilled. Rather, the statistical model simply attempts to identify the portion of unexplained unfulfilled demand that is highly correlated with factors related to counterfeiting and UUPC (rule of law, seizures, etc.). These factors should only influence, in the short term, illegal market behaviour targeting certain markets, not legitimate demand within the markets. Thus, firm inaccuracy or inexperience in forecasting demand will not over-inflate the model's prediction regarding counterfeit. It simply implies that the model will explain less of the unexplained unfulfilled demand (and hence a lower share of counterfeited or UUPC goods) than if firms are good at forecasting (in which case counterfeiting or UUPC will explain a lot more of the unexplained unfulfilled demand).

At this stage, the method makes use of economic theory about the specific structure of a given market in which, for example, several firms are successful in a product market and with that success comes incentive to replicate. We seek to improve the method by gaining insights from

firms on the specific knowledge of their product. We understand that calculating share of future demand (e.g. forecasting) is not an exact science and sectors will have differing influences on their forecasts. There are many uncertainties to calculating demand and often demand patterns may persist- demand this quarter is similar to last quarter or similar to the quarter last year- which can facilitate forecasting. The method takes the unexpected variation in forecasted and actual demand and uses that information to infer information about the level of counterfeiting or UUPC. In order to do this, it would use past information over time to detect a statistical relationship over time between the forecast 'error' and counterfeiting/UUPC. As such, we are seeking past data-quarterly and yearly data in a variety of markets (i.e. data from 2005-2010).

Advantages of this new approach

The approach has important advantages:

1. It is flexible. In the future, it may be possible to use results to calculate other types of impacts such as tax revenue losses, health outcomes, social capital, etc. It also allows for changes in a market and newly identified indicators of counterfeiting/UUPC to improve the accuracy of estimates.

2. It can be tailored but allows for estimation within a unified framework. The method captures characteristics of each sector that differ consistently across sectors. The beta model can be modified and refined to better reflect given sectors and estimate counterfeiting by sector, should that be warranted. At the same time, it applies the same generic approach to all sectors so that estimates of counterfeiting across industries/products could be compared.

3. It provides a direct counterfeit/UUPC link. It does not require supposition about the economic impact that counterfeiting/UUPC has. The beauty of this approach is that, if validated, it should capture counterfeiting/UUPC that occurs from unauthorised producers directly competing in the primary market as well as by stealing sales through a secondary ("knock-off") market.

4. It provides value for money. It meets the criteria of the European Commission to have a method that can reliably be applied to many sectors and countries at a reasonable cost. While superior methods for estimating counterfeiting/UUPC of a specific product exist (e.g. mystery shopping for particular goods, chemical analysis of cigarettes buds or fertilizer), these methods are applied to relatively small samples within a sector and geographic region; they are therefore limited to describing the situation across the EU. Furthermore, these methods are performed at significant cost and extending those methods to every sector and country every year was deemed unfeasible. The proposed method provides value for money by generating a useful, low-cost strategy for understanding trends between periods when more sophisticated methods can be applied.

5. It can be used for policy evaluation. The method allows analysts to investigate the relationship between changes in policy and changes in counterfeiting/UUPC. Should this method prove to generate reasonable estimates, consistent with those generated from alternative methods used by our partners, it will provide a way for the European Commission to measure trends in counterfeiting/UUPC within the EU. It can also directly quantify the effect a policy had in a Member State on the counterfeiting/UUPC market.

Frequently Asked Questions

1. *Exactly what data are required from firms?*

   The method requires firm-specific information on the quarterly (annual or semi-annual) forecasting error of authentic units sold to market. We define this as the difference between the forecasted number of units sold in a given quarter (based on the firm's own forecasting method) and the actual number of units sold in that quarter.

   Brands can be made anonymous by sending 'product A' or SKUs. We only need to know the product category and country for which sales are estimated.

   Ideally, if the firm looks into why forecasts differed from the actual units sold, this information would be passed on by sending an updated forecast error (with this forecast error taking into account these unexpected product or firm occurrences in the market). This aids in the accuracy of the analysis.

2. *How can you guarantee that no sensitive business information will be released to third parties?*

   RAND has extensive experience working with and protecting confidential and private information. In addition to signing confidentiality agreements with participating organisations, standard procedures are in place to guarantee that data are not accidently or inadvertently released in any sort of identifying form. All proprietary data will be stored on encrypted desktop computers that are physically located within a RAND office. RAND offices are secure facilities with guarded access. The buildings themselves have several layers of security due to the national security work conducted here. Security badges are required for entering all locations. Security cameras are in operation throughout RAND's buildings. Non-RAND personnel must always be escorted.

   Computers are physically protected by being located in secured locked offices and then digitally protected through the use of significant firewall protections, network and computer-specific passwords, and encryption. It is a standing rule at RAND that no sensitive data can be stored on a laptop, so as to protect the integrity and confidentiality of the data. De-identified working files are permitted, but these files must be protected via password and laptop encryption. Once a project is completed, all working and original data files are destroyed.

3. *What if firms take into account counterfeiting and UUPC in their quantity forecasts?*

   We need to know if quantity forecasts made by the firm are adjusted to take into account firms' expectations about counterfeit/UUPC, whether directly or indirectly.

   If firms calculate product demand and then reduce it by an amount of counterfeit/pirate, this amount needs to be presented, as well as explanation of where it comes from. Even if it is a constant percentage or amount of total, it is useful for us to understand the previous amount considered and assess whether our model can pick that amount up. We need to know, in particular, whether firms calculated the amount of counterfeiting/UUPC purchased in terms of demand for authentic good only or presumed willingness to buy knock-offs.

4. *What if a firm sells more than forecasted?*

The model is not affected by firms selling more than forecasted, as these are deviations that firms can generally account for post-hoc. Even if firms cannot account for unexpected growth entirely, negative forecasting errors do not influence the integrity of the model. We estimated the model using both absolute values as well as actual deviations and the results were qualitatively similar in terms of predictions from our counterfeiting measures.

5. *What if firms send unexplained forecast errors that are misleading? Can the method be 'gamed'?*

The only way this can affect our model is if all firms within a sector alter their unexplained forecast errors by the exact same amount each year into the future. This would require collusion on the part of firms in a sector. The approach assumes firms do not collude in their submission of unexplained forecast errors. One firm that systematically provides misleading forecasts will not bias the results, as such persistent "errors" in data will be captured statistically through a technique called 'fixed-effect modelling'.

6. *What if factors that contribute to legal sales also contribute to illegal sales?*

It is entirely possible that factors leading to legal sales contribute to illegal sales, and we expect that. Our model is constructed such that it will detect (statistically) a relationship (if there is one) between factors specific to counterfeiting or UUPC, and the forecast error. If the forecast error was caused by market factors that also influence the number of counterfeiting or UUPC, then the model will pick it up, particularly if it is not unique to a given firm (e.g. rise in income means people buy more of all goods, authentic or counterfeit/UUPC). There will be unknown factors we do not include and that is okay for the method. The relationship between these unknown factors and the forecast error ends up in the error term of the model and is left unexplained. Therefore, for example, the model may explain 75% of the forecast error- 5% may go to counterfeiting/UUPC, 50% firm-specific issues (e.g. forecast method), 10% market-specific (e.g. weather), 10% product-specific and 25% is left unexplained and in the error term of our model. This will vary by sector.

7. *What is done about consumers who have purchased a counterfeit product, but because of the much higher legitimate price (or other reasons) would never purchase a legitimate good? How is that counted if firms do not count that in their estimates of counterfeiting/piracy?*

The method does not capture those who would not have bought the authorised product. It captures those who would have bought the legitimate product and were deceived or were subsequently convinced/decided to buy the unauthorised product.

8. *What about counterfeiting sold through the Internet?*

As we are examining firm-level data, rather than breaking down by the multiple channels of distribution, it does not affect our results. We are estimating changes in counterfeiting/UUPC at a producer-level, not at a distribution level in the supply-chain. This means the method does not discern where the counterfeit/UUPC was purchased, only the number of units. Of course, sales on the Internet if not identified by the firm as being sold in a given market, may lead to some error related to specific country sales. That error, if unchanging over the small window of data we are able to secure, will be accounted for statistically using country/market fixed effects.

9. *What if a factor is counted in one country but actually reflects counterfeiting and piracy occurring in another country?*

Nearly all variables we have discussed including in the model refer to activities happening within a country, i.e. national crime rates, rule of law, anti-counterfeiting task force.

The one variable that might be measured in one country but reflect behaviour in another country is seizures, which can be recorded in the country where the good was intercepted, rather than the country of provenance (e.g. destination country). Obviously all the countries with major ports (Rotterdam, Netherlands; Riga, Latvia; Hamburg, Germany) will have higher seizure statistics if this practice happens of attributing to the country where intercepted. In this case, the problem is measurement error. To reduce this measurement error, we would include a variable for seizures in a neighbouring country (as a way of capturing trade being picked up elsewhere).

10. *What are the benefits of participating to this pilot stage to my company?*

The benefit of working with RAND in the pilot stage is twofold: (1) you have more direct influence on how our model develops, as through our discussion with our pilot partners we will be increasing our knowledge both within sectors and across sectors; and (2) you have access to some of the top academic experts knowledgeable on model building and forecasting, which might be useful for the development of your own internal forecasts and/or estimates of counterfeiting. You get immediate use of the model for your own company purposes, and it can be refined and tweaked internally to be made more applicable to your own industry.

11. *Who from my company should be involved in this pilot stage?*

The people who will most likely contribute to the success of this pilot, are those that have most experience with the sales forecasts, its methodology and results. Furthermore, the individuals responsible for identifying, estimating or dealing with counterfeiting and UUPC in your company will have the best understanding of the factors (in the second stage regression) that contribute to counterfeiting and UUPC in your markets.

Appendix C. Road map for implementing additional pilots of the RAND Model

In this appendix we provide a roadmap for future efforts to more fully assess the viability of the RAND approach across products, firms and sectors. It is our hope that our previous work shows sufficient promise that it will encourage firms to partake in a more thorough evaluation of the model, which is clearly warranted before it can be recommended as a new basis for measuring counterfeiting going forward.

The lack of participation from multiple industry sectors and product types is a real weakness of the pilot work conducted thus far. Without a proper evaluation of sufficiently varied products, firms and industries, it will not be possible to assess to what extent variables included in the second stage of the model should be expanded or refined when estimating counterfeiting within particular industries or markets. Moreover, it may be the case that the model performs better tracking counterfeiting in some industries and worse in others. Thus a more thorough evaluation of the approach is warranted and encouraged.

While limitations may be found, it should be noted that the model (as it has been conceived) is flexible enough to accommodate many firm, product or industry specific factors that might be unique. Statistically, much of these unique factors (even if they are unobserved) can be included in the model through fixed or random effect estimation techniques once more data have been collected. Indeed the current model included product-specific fixed effects which improved the predictive power of the model with the data we have, and thus it was not necessary to model every time-invariant factor that might be important for predicting deviations from forecasts.

1. Selection of industry partners

The generalizability of findings from any pilot evaluation depends critically on the validity and representativeness of the data used to test the model. From a statistical perspective, the ideal sample on which to test this methodology would be a truly "random sample" of firms, products and sectors such that we are able to represent the types of products and industries that are targeted by counterfeiting. However, it is rarely possible to generate such a sample. As the purpose of this exercise is to assess the utility of a model to accurately measure trends in the amount of counterfeiting occurring in a regional market, one would probably not want to start with a random sample of all firms anyway, but rather a list of firms stratified by the likelihood that the firm's product is heavily targeted by counterfeiters. Thus, you would want to draw more firms from those that are likely to be big targets and fewer firms from those who are less likely to be targeted or targeted at a low level. An additional advantage of drawing firms in a stratified manner is that many firms that are frequently targeted by counterfeiters often already have in place their own process for estimating the

amount of counterfeiting affecting their own sales, and hence an independent measure is available that can be used to validate trends identified by the RAND model for that firm.

Implementation of a truly stratified random sample is highly unlikely given that it requires participation by the firms and release of somewhat sensitive data. Given that a selective sample is the most likely outcome, we suggest the following considerations when developing a sample of firms/products:

1. Approach firms who have a high concentration ratio and/or brand value such that they are able to secure above-normal economic profits. Typically these are firms that have patented products and/or well branded names and usually represent the same products/firms heavily targeted by counterfeiters. If the model does reasonably well identifying trends based on data from firms heavily targeted by counterfeiters, than this will represent the bulk of the market policymakers and individual firms are interested in understanding.

2. Target firms in each of the following five industries: Apparel, Footwear and Designer Products (Textiles); Automotive; Consumer Electronics; Food and Beverage; and Pharmaceutical Industry. Obviously, more industry representation would also be good and should be encouraged, but these five industries contain products sold by 18 of the top 20 branded values in 2006 according to data from Interbrand and are industries well understood to be targeted based on seizure data. Moreover, the industries represent a range of products with fairly different regulatory structures (and hence difficulty to break in/sell their products); production and distribution processes; consumer behaviour (repeat customers of a frequent purchase versus occasional customers of infrequent purchases), and products with varying health/safety risks to consumers. Inclusion of firms within each of these industries, therefore, will provide a range of issues that must be dealt across important dimensions, including the nature of the product, consumer, or supply chain. Thus, such firms will provide the most insight in terms of how to build a sufficiently broad model to handle these issues.

3. Work with firms that already have in place forecasting technology. While a few small firm representatives suggested that they did not generate internal forecasts, we found this to be a rare event among international companies who have to figure out in advance how much product to ship to specific markets to meet expected demand. As it is costly to move products and often cannot be done cheaply on short notice, plans get put in place regarding how much product will go where. This is the basis of a forecast for any firm. Many firms do in fact do this – under the guise of "logistics" - but not always in a formalized structure. If the goal is to quickly assess the reasonableness of the approach, then working with firms that already do this will facilitate a timely test of the model.

4. Work with firms that have retained historical information on forecasts and actuals. A major limitation of any model relying on indicators of counterfeiting measured at the country level is that these data take time to get produced and released from the various agencies collecting them (we observed at least a one year lag on most of the measures we were interested in examining). As such, it is important to identify firms who can generate information on previous forecasts and actuals going back in time, at least for a few years, in multiple country-level markets. Doing so will generate more statistical power and variability (over time and across countries), which will improve the precision of the model and generate a better test to reliability. However, it is important *not* to rely on data just from 2008-2010, due to the tremendous uncertainty caused by the global Great Recession. Any model, even one proven to be valid and reliable, would

likely generate inaccurate results of trends if based solely on data from this time period given the unusual circumstances firms were operating in.

Ultimately the final decision of which firms to include in a full blown pilot will depend on the willingness of selected firms to engage in this research jointly with OHIM or another agency executing the pilot. The lack of a truly random stratified sample should not discourage the piloting of the method, because much of the concerns of the model relate to the technical implementation of it, which can be demonstrated regardless of which firms participate. To the extent that participating firms have their own estimates of counterfeiting using their own preferred methods, then estimates from the RAND model can be compared analytically to those firm estimates to assess the quality of the model. It is expected that the level of counterfeiting in any given market/year may be off by a relatively large factor depending on the quality of the RAND model.

2. Collection of relevant data

Two different types of data are required for estimating the model. First, firm specific information on unexplained forecasting errors for various products targeted by counterfeiters is needed. Second, data collected from national government and international agencies reflecting factors that might influence the level of counterfeiting within a national market are required. We will discuss each of these in greater detail here.

2.A. Firm specific information on unexplained forecasting errors

Most firms have to decide how much product to (1) produce, and (2) ship to a given market far in advance of the sale actually occurring, as most businesses understand that it costs money to overproduce (due to having to store excess product and/or dump it at lower prices) or under-produce (as more goods could have been sold if shipped to Market B instead of sitting at Market A). Thus, on some level multi-national firms have to do forecasting of expected sales in particular markets in order to maximize the profit to be gained in each or they lose money. The level and detail of that forecasting can differ tremendously, but to some extent they all boil down to an expectation of the number of units (or dollar revenue that can be translated into units) sold in each market, which often get based to some extent on information regarding units sold in the past. Those firms that are owned publicly and are sold on national or international stock exchanges regularly produce projected earnings based on expected units sold of given products so as to encourage investors to put their money into their company. When they do not meet their projected earnings, then they have to explain to the same investors why they did not meet their forecast and/or how they exceeded it.

Thus, it is not unreasonable to presume that large, public international firms do in fact collect data systematically on forecasted units sold of various products sold and conduct regular look-backs to assess why their projected sales were off from actual units sold. Some differences can be explained by firm specific factors (issues in the supply chain, rising distribution cost, production problem), competitor behaviour, or larger macro factor (political factors, national recession, major national disaster). To the extent that a firm can say, "Well x% of our error was due to things we now understand to have thrown us off," then the analyst knows that 100-x% can be considered "unexplained forecasting error." Determination of that percent that is unexplained would ideally be done systematically, as we did in our work with the industry partner in the pilot using their firm level data in a regression to ascertain the forecasting error that could not be explained by firm, competitor, and market factors. In that way differences in unexplained forecasting errors across national

markets could be accurately captured. If, however, a firm person can provide an informed estimate of this information by national market, that should suffice in terms of the needs of the first stage regression provided that the unexplained forecasting error estimated by the analyst adjusts reasonably well for factors that become realized by the firm and are not inclusive of expected counterfeiting. If expected counterfeiting is included, then that must be approximated itself and added back into the "unexplained" component.

Based on our experience talking to firms, we have put together a template (Exhibit A) that can be used by OHIM (or another central agency responsible for conducting future work in this area) to assess the extent to which potential firm partners have the data necessary to estimate the RAND two stage model. The first few questions assess the extent to which the firm collects forecasting data, over what period of time and for what national markets, and in what units. While units do not have to be identical across products, it is important to make sure that some products are not estimated in very small units while others are measured in large units, as it will cause the model parameters to be too highly variable and possibly inaccurate. It would be better to translate all firm data to a broadly consistent measurement, such as market level retail or wholesale level categories. Making sure all data for a given industry are represented at the same market level ensures that the model generates reasonable estimates.

Exhibit A: Firm Questionnaire

1. What are the units in which data is collected (forecasts & actual) on this product? (i.e. tons, litres, packs, units sold)

2. Our goal is to obtain annualized nation-level historical data for a minimum of five years (preferably longer). How many years of forecasted data might you be able to retrieve (in terms of units)?

3. At what frequency do you forecast and collect actual units sold? Circle One.

    Monthly   Quarterly   Annually

4. Do you (or can you) generate forecasts at the National or Member State-level?
    Yes / No

    4.1 If no, at what level and for which countries do you have data?
    Level of market analyzed

    _____

    Countries represented

    _____

    _____

    _____

5. Once actual unit sales data are known for your market period specified in #3 above, do you go back and analyze your previous forecasts to determine how

much of the forecast error could be explained (i.e. the difference between the # of units you forecasted to sell of this product and the actual units sold by market)?

Yes   ➜   Go to Question 6

No   ➜   Go to Question 9

6. When you analyze your forecast error, which of the following do you consider is a possible explanation of the forecast error (check all that apply):

___product-specific supply chain, inventory or marketing issues
___company-specific changes or issues
___unexpected product releases or marketing campaign from competitors
___unexpected market factors (supply shocks, demand shocks)
___unexpected policy changes in a targeted or transport market
___ counterfeit production
___ Other _____

7. Do you retain estimates of how much of the forecast error is explained (e.g. 80%, 100%, 20%) by period going back five years or more (or can you reconstruct these off of your forecast analysis)?

Yes / No

8. Would you be able to send us information of the following type?

Yes/ No

Product	Time	Country	Forecast error	Units	% Explained Post Actuals
A	2001q1	Austria	-1000	Tons	80%
A	2001q2	Austria	1000	Tons	45%
A	2001q3	Austria	0	Tons	100%
.	(continues for more time periods)		.		
	2001q1	Belgium	100	Litres	88%
A	2001q2	Belgium	-101	Litres	97%
A	2001q3	Belgium	50	Litres	100%
.	(continues for more time periods)		.		

9. If you do not go back and examine forecast errors, do you have a sense of what sort of factors typically drive your forecast from being off?

___product-specific supply chain, inventory or marketing issues
___company-specific changes or issues
___unexpected product releases or marketing campaign from competitors
___unexpected market factors (supply shocks, demand shocks)
___unexpected policy changes in a targeted or transport market

___counterfeit production
___Other _____
___I do not know

10. Would you feel comfortable attributing a certain fraction of your forecasting error to each of the factors above? If yes, please specify what proportion is driving each.
___product-specific supply chain, inventory or marketing issues
___company-specific changes or issues
___unexpected product releases or marketing campaign from competitors
___unexpected market factors (supply shocks, demand shocks)
___unexpected policy changes in a targeted or transport market
___counterfeit production
___Other _____
___I do not know

The fifth question in Exhibit A indicates to the outside analyst the extent to which the firm evaluates the previous forecast after the period has passed, as is expected in the first stage of the RAND model. If the firm analyst responds in the positive, the sixth question tries to identify what sort of factors are being included, and explicitly identifies whether expectations regarding counterfeiting is something that is built into either the forecast or the forecasting error. If the firm does not systematically evaluate factors driving the firm's forecast error, the respondent is queried in questions 9 and 10 as to whether they do any informal assessment of why actuals are off from forecasts. If neither is done, then the RAND two stage model could not be applied to data from this firm, and instead the simplified model must be used.

Question 8 in Exhibit A provides an example of the ideal data that would be collected from firms. It is our experience; however, that it takes a bit of back-and-forth conversation with the firm analyst before the data as represented in question 8 can be meaningfully produced in a way that is useful to the outside analyst. Issues of how to aggregate information across product groups and/or markets to get to country level data may come up, but are easily worked through with a little bit of direct communication the initial time the data is being put together. An understanding of firm specifics in their estimation process is useful for ensuring that what gets sent to the analyst and called "unexplained forecast errors" are as clean as possible. Once the initial time investment is made with each firm, updates become a relatively easy and seamless process.

To the extent that firms understand the objective of the first stage and are willing to provide data on unexpected forecasting error to a centralized agency, then the first stage model does not in fact have to be estimated by an outside or centralized analyst. Instead, firms could simply provide to the outside analyst the following information:

- an unidentified product number (the firm would be the only entity that would know what the product was) – multiple products from each firm may be targeted and could be submitted using different de-identified product numbers
- the total forecasting error generated for a given market for a fixed period (quarter or year)
- the percent of that total forecasting error (in units of product) that remains "unexplained" after realized values of important firm, competitor, or market factors are taken into account.

- For modelling purposes, information regarding which industry the product is sold would be helpful (so models could be run separately by industries; specific categorical industries can be predetermined so uniformly applied) as well as a firm dummy variable (so that data can be updated to existing firm data without actual identification of firms by the analyst).
- To facilitate translating units to monetary values post estimation, an average price per unit (for the market level reflected by the forecast) would also be needed. With that information, de-identified in terms of specific products, the outside analyst has everything s/he needs.

Alternatively, if the firm is not able or prepared to estimate the first stage itself, then the relevant data could be sent to the central outside analyst, as was done in the RAND pilot of the model. This, however, requires release of additional sensitive data, which require special data use agreement and arrangements to protect all firm sensitive information. It is presumed that firms would be less willing to send sensitive information directly to a centralized agent, but might be willing to work with a third party to generate estimates from a model similar to that estimated by RAND in this pilot.

2.B. National and international data on factors influencing counterfeiting

For the purposes of our pilot study, we collected numerous variables that the literature suggested could be correlated with counterfeiting. However, some indicators did not vary much over the relatively short period of time that we were analyzing data, and other indicators were not consistently reported for all countries in which we had product data. Thus, we do not believe that the variables included in our second stage regression model are the only counterfeiting variables that should be used in a model using data from other industries with more years of data. Exploration of a range of additional counterfeiting variables can and should be done, based on the variation that is available across markets, across products and across time. If industry and country fixed effects are included in all models, then factors that do not vary over time will be dropped out of the model empirically. That does not mean that these variables are not important, but rather that their influence is being captured by the fixed effect.

Through the literature review, we identified a number of potential indicators related to counterfeiting and piracy that might be worth considering in future pilots, only some of which could be included in our analysis. As discussed in prior reports, we group the set of indicators into two broad categories:

- Country-specific institutional, economic and social indicators that are likely to affect counterfeiting and piracy rates across a wide-spectrum of products; and
- Country-level technological indicators that will have ramifications specifically for products that are affected by access to and utilisation of technology.

Country-specific institutional, economic and social indicators. Institutional indicators capture political and legal frameworks that are likely to affect counterfeiting and piracy rates within a country. Three indices used in the RAND pilot to measure the strength of a country's legal and political frameworks were the rule of law, government stability and the control of corruption indices. Economic indicators, which capture both individual demand for C&P goods and potential distribution channels by which counterfeit products can enter a country's product markets, were also included such as the log (per capita income), GDP growth, and measures of international trade. Income inequality is another measure that has been shown to be important in the empirical research, but did not sufficiently vary in the period we analysed to be included in our

analysis. Similarly, social indicators that capture the influence of collective or shared perception of consuming C&P products, such as attitudes or perceptions of social consequences of purchasing fake products, could also be included if there is sufficient variability over time in the indicators. Examples for all these metrics are included in Table C1.

Information and Communication Technology Indicators. The second set of indicators that are hypothesized to be important for understanding the unexplained variation in the second stage of our empirical model are related to the ability of individuals in a country to access and utilise information and communication technologies (ICTs). While obviously important for pirated goods, they are also likely to be important for goods that are sold over the Internet. The International Telecommunications Union has developed an index to capture the availability of ICTs within a country including measures of fixed telephone lines, mobile cellular telephone subscriptions, international Internet bandwidth, proportion of households with computers and proportion of households with Internet access at home. The ICT utilisation variables include measures of the number of Internet users, fixed broadband Internet subscribers and mobile broadband subscriptions. ICT skills are measured through the adult literacy rate, secondary gross enrolment ratio and the tertiary gross enrolment ratio.

Table C8-1: Potential measures and data sources for 2nd stage regression

Measure	Category	Data Source and Years
Rule of Law index	Institutional Factor-Crime and justice criminal	World Bank: Rule of law indicator, 1996-2002 (Biennial), 2003- current (Annually)
Control of corruption index	Institutional factor-Crime and criminal justice and Political Environment	World Bank: Control of corruption indicator, 1996-2002 (Biennial), 2003- current (Annually)
Government stability index	Institutional Factor-Political environment	World Bank: Political stability indicator, 1996-2002 (Biennial), 2003- current (Annually); World Bank (International Finance Corporation) -World business environment survey, 2000; Transparency International: Corruption perception indices, 2001-2010
Seizures	Institutional Factor-Crime and criminal justice	European Commission (Taxation and Customs Union): DG TAXUD, 2003-Current
Number of annual infringement cases (copyright, patent, trademark)	Institutional Factor-Crime and criminal justice	European Commission (Taxation and Customs Union): DG TAXUD, 2003-Current
Presence of an anti-counterfeit/piracy task force	Institutional Factor-Crime and criminal justice	European Commission
Per-Capita Income	Economic - Income	World Bank
Income inequality	Economic - Income	World Bank, 1981-current
Economic growth	Economic - Income	European Commission (Eurostat)
Labour force participation	Economic - Income	European Commission (Eurostat)
Inflation	Economic - Income	European Commission (Eurostat)
Sales tax rate	Economic - Income	TMF Group
Market share of the	Economic-Profitability	Various

Measure	Category	Data Source and Years
legitimate product		
Level of trade with other countries/trade dependency	Economic- Distribution Channel	IMF; European Commission
Value added from tourism	Economic - Other	European Commission (Eurostat)
Proportion of population speaking English	Economic - Other	Eurobarometer Surveys, I973-2010
Openness to innovation	Social - Willingness to buy	Various
Brand loyalty	Social - Willingness to buy	Various
Anti-counterfeiting campaigns	Social - Willingness to buy	Various
ICT Development Index	Technology	United Nations (ITU) – 2002-2010

3. Estimation of the model

As stated previously, the first stage model would ideally not be estimated by a central analyst as firms are in a better position to generate estimates of unexplained forecasting error than a central analyst would. However, if firms are willing to provide the necessary information identified in Exhibit A allowing for estimation of a model of forecasting error (in a de-identified fashion so as to protect each firm's identity), a central analyst could estimate the first stage model if doing so places less burden on individual firms. The specific formulation of that model will depend on the data made available by each firm.

A more likely scenario, however, is one in which firms provide to the central analyst market/year estimates of the unexplained forecasting error, having already netted out product, firm and market shocks that were observed post-hoc by the firm. Assuming that this is the case, than the proper specification of the second stage model in terms of included counterfeiting variables and pooled data will ultimately depend on statistical tests of the empirical model. This in turn, will also be influenced by the specific national markets for which data is available on the products (more countries will reduce the potential collinearity) and the number of years for which data is available (a shorter time period will increase the probable collinearity). It may be the case that the second stage model has greater predictive power when estimated separately by industry, so that specific sets of variables (which are less collinear for that given industry) can be used to predict the number of units of counterfeiting of products within that industry. The sum of predicted counterfeiting units generated by these models for each product included within an industry can then be added to units estimated from other industry models.[44] Trends in the total number of units counterfeited should be analyzed in terms of units sold rather than monetary value, as trends in monetary value could move due to changes in the average price of products sold rather than changes in the number of units produced. Nonetheless, estimates of the value of units sold might be useful for representing the total value of

---

[44] Assuming there is sufficient statistical power, a fully interacted model which interacts an industry dummy variable with all the other right hand side variables included in the model could be tested through an F-test of the joint significant of all the interacted terms. If the interaction terms are jointly significant, then separate models by industry are more appropriate.

counterfeits produced in a given year (and by indexing changes in average price of units sold might assist in differential trends in values from units sold).

It will take careful statistical analysis of the second stage model to ultimately determine the most appropriate specification for predicting counterfeiting. While the previous discussion named a variety of country-specific variables that might be used to capture the amount of counterfeiting in a given market, it is unclear which variables will have independent variation over the time period that will be analyzed or the unique markets for which firms can consistently provide data. To the extent that fixed product, firm, industry and country effects can help improve the precision of the model and improve estimates, these should be considered. It is possible that including country-fixed effects, however, might completely eliminate the predictive power of counterfeiting variables if the time period in which data in available is too short, and hence careful consideration of the within country variation in counterfeiting variables over time will be important for understanding the benefit of including country-fixed effects in the second stage model.

Additional data from more firms, more products, and more industries will provide useful insights on the best specification of the model and the key factors that really drive that model. Once the best specification is determined, it can be used predictively with expectations of values of counterfeiting measures in particular markets to predict how changes in particular variables can influence the trend in counterfeiting within a given national market or regional market.