Institutional and Empirical Essays on Regulatory Impact Assessment

Thesis submitted for the Degree of Doctor of Philosophy

2006

by Seán Lyons

University of Dublin, Trinity College

Declaration

This thesis has not been submitted as an exercise for a degree at this or any other university. It is entirely my own work with the partial exception of Chapter 6, which includes work co-authored by others that has been accepted for publication. Such work is acknowledged in the text. I agree that the Library may lend or copy this thesis upon request.

_____ _____

Author's signature Date

Summary

This thesis is concerned with institutional arrangements and empirical tools that may be used to assess the impact of regulation on the economy, and thereby improve regulatory policymaking.

We start in Chapter 2 with a qualitative introduction to regulatory impact assessment, highlighting some of the challenges facing small regional governments wishing to improve the quality of regulation they enact.

The remainder of the thesis is divided into two parts. Part 1 considers and evaluates the optimal design of tests aimed at discovering specific classes of regulatory impact. Our particular example is a negative clearance test for the effect of regulations on competition. One such tool, the UK competition filter, has been in operation in the UK since 2002. Applying it to case studies from Ireland (in Chapter 3), we show that particular design features may easily lead to excessive false negative results. We then propose a more appropriate structure for tests of this kind. A paper based on this chapter is forthcoming in the *Journal of the Statistical and Social Inquiry Society of Ireland*. The analysis is extended in Chapter 4 to show that the UK filter is also prone to give excessive false negatives in its own jurisdiction; a paper drawing on these results and co-authored with my supervisor, Dr. Francis O'Toole, is under review by the *Journal of Competition Law and Economics*.

Part 2 focuses on empirical studies of regulatory impact. Two of the three chapters in this part are concerned with aspects of mobile telephony regulation. The first of these, Chapter 5, notes that an increasing number of countries require mobile telephone networks to offer mobile number portability (MNP), which allows customers who wish to switch mobile operator to keep their mobile numbers (including the prefix) and avoid the costs of switching to new numbers. *Ex ante* assessments suggest that MNP should reduce switching costs and strengthen competition. We construct a new international dataset that allows us to estimate econometric models of the benefits of MNP and we test MNP's impact on market outcomes using this international time-series cross-section data. The results show that MNP reduces average prices and encourages churn (a proxy for switching) when the switching process is rapid (e.g. less than 5 days) but not when it is

slower. A paper based on the research in Chapter 5 is under review by *Information Economics and Policy*.

Chapter 6 also employs the mobile telephony dataset introduced in Chapter 5, focusing this time on a metric popular among regulators and market analysts: average revenue per user, or ARPU. This chapter, which draws on a co-authored paper forthcoming in *Telecommunications Policy*, develops an econometric model of the determinants of ARPU and shows that this statistic may not be a good indicator of competitive conditions. ARPU proves to be better explained by subscribers' incomes than by market concentration measures.

In Chapter 7, we ask how far the constraints perceived and reported in surveys of small businesses (e.g. excessive regulation, access to capital) are reflected in such firms' subsequent survival rates. There appears to be a relationship, but it is neither as strong nor as straightforward as the popularity of this sort of evidence among policymakers might lead us to expect. Indeed, the perceived barriers expected to be most significant on the basis of previous research – excessive regulation and constrained access to finance – show no significant relationship to survival rates. Until more convincing evidence is found showing that this type of survey evidence helps explain market outcomes, using it in the assessment or design of public policy may be problematical. The chapter suggests ways of enhancing surveys so that the real effects of perceived constraints can be more clearly identified. A paper drawing on it is currently under review by *Small Business Economics.*

Acknowledgments

I am particularly grateful to Francis O'Toole, my supervisor, who provided invaluable advice and comments on the thesis as a whole and on various drafts of all chapters. Patrick McCloughan of Indecon co-authored the paper on which Chapter 6 is based and commented in detail on Chapters 3 and 7. Tommaso Valletti and two anonymous referees at *Information Economics and Policy* provided extensive and helpful comments on a paper based on Chapter 5.

Participants at Trinity College seminars gave suggestions on early results of the research in Chapters 3 and 5. A paper based on Chapter 3 also received many useful comments when I presented it to the Statistical and Social Inquiry Society of Ireland, especially from the discussants, Cathal Guiomard and Paul Goreki. Chapter 2 benefited from class discussions during my lectures on the Masters in Economic Policy Studies at Trinity and comments by seminar participants at the Policy Institute. Thanks are due to Frances Ruane for the opportunity to present my work in both of those fora and for discussions on the relationship between regulatory impact assessment and evaluation of public expenditure.

I am grateful to Elaine Pryor at Merrill Lynch, who kindly allowed use of the Merrill Lynch data referred to in Chapters 5 and 6, and to the European Commission and EIM (especially Jan de Kok) for providing access to the ENSR dataset discussed in Chapter 7. Participants at the 2006 INFER Conference in Cork, particularly John Sell, contributed many helpful comments on the research discussed in Chapter 7.

Finally, thanks are due to my wife Mary and daughters Clodagh and Nessa who never complained about the hours I spent on this research.

Of course, any remaining errors are my own.

Table of Contents

Tables and Figures

1 Introduction

"The problem of cost-benefit analysis is simply whether we can find workable shortcuts," suggest Atkinson and Stiglitz (1980, p.475). Even where expected aggregate costs and benefits of a project are known, it may not be straightforward to arrive at total welfare effects.[1] One might add that the ideal data are never available and analytical time is costly. If this characterisation has resonance when discussing public expenditure evaluation, it seems even more appropriate for the *ex ante* analysis of regulatory measures: regulatory impact assessment (RIA). At least the expected cost to the exchequer of proposed public expenditures tends to be known in advance: for prospective regulatory measures, even the *sources* of likely costs and benefits may not be known to the policymaker.

In this thesis, we discuss and implement some workable shortcuts. Our aims are twofold: to contribute to the empirical literature by measuring the effects of various regulatory measures and to provide practical advice for policymakers involved in the evaluation of regulatory impact, particularly those in small regional jurisdictions such as Ireland.

We start in Chapter 2 with a qualitative introduction to regulatory impact assessment, highlighting some of the challenges facing small regional countries wishing to improve the quality of regulation they enact.

The remainder of the thesis is divided into two parts. Part 1 considers and evaluates the optimal design of tests aimed at discovering specific classes of regulatory impact. Our particular example is a negative clearance test for the effect of regulations on competition.

[1] For example, assessments generally require explicit or implicit distributional judgements.

One such tool, the UK competition filter, has been in operation in the UK since 2002. Applying it to case studies from Ireland (in Chapter 3), we show that particular design features may easily lead to excessive false negative results. We then propose a more appropriate structure for tests of this kind. A paper based on this chapter is forthcoming in the *Journal of the Statistical and Social Inquiry Society of Ireland*. The analysis is extended in Chapter 4 to show that the UK filter is also prone to give excessive false negatives in its own jurisdiction; a paper drawing on these results and co-authored with my supervisor, Dr. Francis O'Toole, is under review by the *Journal of Competition Law and Economics*.

Part 2 focuses on empirical studies of regulatory impact. Two of the three chapters in this part are concerned with aspects of mobile telephony regulation. The first of these, Chapter 5, notes that an increasing number of countries require mobile telephone networks to offer mobile number portability (MNP), which allows customers who wish to switch mobile operator to keep their mobile numbers (including the prefix) and avoid the costs of switching to new numbers. *Ex ante* assessments suggest that MNP should reduce switching costs and strengthen competition. We construct a new international dataset that allows us to estimate econometric models of the benefits of MNP and we test MNP's impact on market outcomes using this international time-series cross-section data. The results show that MNP reduces average prices and encourages churn (a proxy for switching) when the switching process is rapid (e.g. less than 5 days) but not when it is

slower. A paper based on the research in Chapter 5 is under review by *Information Economics and Policy*.[2]

Chapter 6 also employs the mobile telephony dataset constructed in Chapter 5, focusing this time on a metric popular among regulators and market analysts: average revenue per user, or ARPU. This chapter, which draws on a co-authored paper forthcoming in *Telecommunications Policy*, develops an econometric model of the determinants of ARPU and shows that this statistic may not be a good indicator of competitive conditions. ARPU proves to be better explained by subscribers' incomes than by market concentration measures.

In Chapter 7, we ask how far the constraints perceived and reported in surveys of small businesses (e.g. excessive regulation, access to capital) are reflected in such firms' subsequent survival rates. There appears to be a relationship, but it is neither as strong nor as straightforward as the popularity of this sort of evidence among policymakers might lead us to expect. Indeed, the perceived barriers expected to be most significant on the basis of previous research – excessive regulation and constrained access to finance – show no significant relationship to survival rates. The chapter suggests ways of enhancing surveys so that the real effects of constraints can be more clearly identified. A paper drawing on it is currently under review by *Small Business Economics.*

At the end of the thesis, a full bibliography and index are provided.

---

[2] As noted at the end of Chapter 5, another empirical paper is in preparation on mobile market share dynamics using an extended version of the Chapter 5 dataset; the intention is to submit this to an industrial economics journal.

2 Regulatory Impact Assessment in a Small Regional Economy

In this chapter, we define regulation and note its growth in comparison to other forms of government intervention in the economy. We then identify some of the institutional options available for improving the quality of regulation and discuss some of the challenges faced by small regional jurisdictions wishing to employ these options. This chapter is intended primarily as background and motivation for the institutional and empirical topics discussed in the chapters that follow.

2.1 Defining "regulation"

The "regulation" to which we refer in this thesis is defined broadly, encompassing not just the actions of economic regulatory agencies but other sorts of economic, social and administrative measures imposed by government. Our use of the term can perhaps be best explained with reference to Spulber's (1989, p.37) definition:

> "Regulations are general rules or specific actions imposed by administrative agencies that interfere directly with the market allocation mechanism or indirectly by altering consumer and firm demand and supply decisions."

Like Spulber, we are primarily interested in the effects of regulatory measures on agents in the economy rather than the legal form of the instruments used to impose it. However, there are two important differences between Spulber's definition and the scope of regulation examined in this thesis. First, we do not restrict ourselves to regulation imposed by administrative agencies; regulation is of interest regardless of the level of government (e.g. national legislature, executive agency, local body) that applies it.

Second, regulatory effects that work through the demand side are largely outside our focus. While the normal scope of regulatory impact assessment extends to actions that

affect the market allocation mechanism (e.g. price controls) and those that affect supply decisions (e.g. health and safety regulation), it tends not to cover important classes of measures affecting demand decisions. The latter, particularly taxation, public expenditure and public investment measures, are generally handled through parallel sets of evaluation arrangements.

This is a somewhat artificial distinction, inasmuch as measures altering demand are often seen as substitutes for those affecting supply or allocation mechanisms; for example, Ireland's "plastic bag tax" was seen as an alternative to command-and-control regulation as a way of reducing litter. However, we may justify focusing on analysis of regulatory interventions on the basis that these are invariably subject to less stringent evaluation regimes than government tax or spending measures.

Nevertheless, it is important to remember that regulation is a form of government intervention, analogous to, and often substitutable with public spending programmes, government investments, subsidies and tax incentives.

2.2 The growth of regulation

In Ireland, and indeed in Europe generally, the use of statutory regulation has risen in recent decades when compared to other forms of government intervention. This is partly due to the shifting borders between public and private activity as privatisation and liberalisation have replaced the earlier European penchant for state ownership and control in many sectors.[3] There may also be an intrinsic bias towards use of regulatory mandates instead of other measures due to policymakers' incentives. In particular, governments

[3] Majone (1996), p.48.

invariably place tighter institutional controls on public programmes and projects funded by the exchequer (e.g. through budgetary processes and *ex post* review requirements) than they do on proposed regulatory measures.[4]

The net effect of these structural developments and incentive effects has been a long-term rise in the usage of statutory regulation in comparison to direct public expenditure programmes.

Given the importance of regulation in modern economies, many countries have put in place institutions designed to ensure that regulatory measures are well designed before they are enacted. In the next sub-section, we discuss some of the options put forward or actually implemented for improving such measures.

2.3 Institutional measures for improving the quality of regulation

Like other forms of government action, regulation may or may not be optimally applied in terms of its effect on societal welfare. There are several schools of thought on the role of regulation, and each suggests different ways in which regulation may go wrong and thus different roles for measures to improve regulatory outcomes.

Up to the 1960s, the standard economic account of regulation held that it was an institutional tool through which governments addressed market failures. This view is sometimes termed the "public interest" school or "normative analysis as a positive theory" (NPT).[5] In an extension to this approach, Demsetz (1969) considered that regulators should not attempt to correct all supposed market failures, but rather should select the

[4] DeMuth (1980a), p.15.

[5] Joskow and Noll (1981), pp.35-36.

policy best able to correct the problem that has been identified. In some cases, lack of a proportionate, efficient instrument might mean that the right choice is to refrain from intervention, even where a genuine economic problem has been identified. In other cases, it may be possible to identify changes in the form of regulation that will improve its effectiveness and minimise the attendant costs. For example, the use of incentive-based regulation rather than command-and-control measures has been extensively studied and has begun to have an impact on policy choices in many countries.[6]

Regulation may still be imperfect under NPT, if only as a result of errors due to imperfect information. In this school of thought, improving regulation would be a matter of improving information about costs and benefits of intervention and better understanding how, or whether, available remedies can address market failures.

For an illustration, consider the development of standards limiting the lead content of petrol, which EPA (1987, Ch.4, p.6) argues were significantly improved through detailed cost-benefit analysis in the mid 1980s:

> "This analysis revealed that reducing the lead content in gasoline from 1.1 to 0.1 grams per gallon would reduce adverse health effects and medical care and educational costs for children with high blood lead levels; could reduce deaths, illnesses, and lost wages from cardiovascular and other diseases; would reduce emissions of other pollutants; and would improve fuel economy and reduce motor vehicle maintenance costs. The present value of the net benefits to the nation from

[6] The increasing use of environmental taxes and tradeable permits in many jurisdictions is one example, see e.g. Hahn (1989). Another, perhaps with wider international acceptance, is the application since the 1980s of retail price regulation using CPI-X price caps to correct the poor incentive properties of "rate of return" and "cost-plus" regulation such as those identified by Averch and Johnson (1962).

> 1985 through 1992 of lowering the lead standard to 0.1 grams was calculated to be $6.7 billion…In large part because the benefit-cost analysis showed such dramatic net benefits, EPA revised its lead in gasoline standard in 1985 to 0.1 grams per gallon…Although EPA might have adopted this revision even without the benefit-cost analysis, that analysis, and the increase in net benefits it showed, provided a strong justification for the revision."

However, in contrast to the NPT, a significant strand of research since the early 1970s suggests that much regulation is not truly intended to improve the welfare of society. Stigler (1971) and the "regulatory capture" literature that followed him sees regulation as primarily a device through which interest groups compete in a political marketplace to capture rents from the public. To the extent that this is the case, regulation is a threat to welfare rather than a possible means of improving it.

The regulatory capture literature argues that improving outcomes is not simply about improving the information possessed by regulators; welfare improvements can be gained by preventing, reducing or eliminating regulation in sectors where it is prone to reduce competition. Several examples of actual restrictions that might fall under this heading are discussed in Chapters 3 and 4, including statutory limits on market entry by taxi drivers, pharmacists, dental hygienists and interstate trucking firms.

More recent studies[7] have suggested that both institutions and markets affect the optimal design of regulation, and both welfare improving and capture-type regulation may be enacted depending upon the institutional setting. The notion that a regulator is an agent in

[7] Peltzman (1981), p.372, coined the term "creeping realism" for this eclectic school, but it does not seem to have stuck. Levine and Forrence (1990) attempted a formal synthesis.

a wider political and economic game, together with parallel work on informational constraints and transaction costs faced by regulators has led to a rich theoretical literature (e.g. Laffont and Tirole, 1993). Issues such as informational asymmetries between policymakers and regulated entities, the need for coalition-building to underpin support for reform, and the overall issue of principal-agent problems within government are emphasised.[8]

While there are disparate views on the rationale for regulation in specific cases, there is a broad consensus that it is not always optimally applied in practice.

On foot of these developments, many countries have put in place institutional measures to improve the quality of regulation. In the next part of this chapter, we list some of the institutional mechanisms used or proposed for this purpose and highlight key policy choices that arise for a country wishing to assess regulation.

Mechanisms for improving regulatory quality can be classified into those with mainly *ex ante* or *ex post* effects.

Ex ante mechanisms

- Regulatory impact assessment, a family of formal quality control processes applied to proposed regulatory measures; see OECD (1997) and EU Presidencies (2004) for examples;
- Requirements for regulatory "budgeting," whereby quantitative constraints of some kind (usually restricted to some measure of regulatory costs) are placed on

[8] Walsh (2004), pp.5-6.

enactments made in a given period, first suggested by Robert W. Crandall[9] and developed in DeMuth (1980b);

- Targets for periodic reduction of regulatory burdens, e.g. by percentage of costs affected or by requiring matching reductions when new measures are adopted ("compensatory simplification"); and
- In the case of independent regulatory agencies, indirect pressure may be exerted by other arms of government towards some preferred regulatory outcome, for example by varying regulatory agency resources such as budgets or staffing.[10]

Ex post mechanisms

- Cost-benefit assessments of existing measures;
- Aggregate (and often more partial) analyses, such as quantification of existing regulatory burdens by applying a uniform measurement methodology such as the "Standard Cost Model"[11] or by conducting meta-analyses of agency estimates;[12] and
- External oversight of regulatory decisions, e.g. via courts, appeal boards or parliamentary committees.

---

[9] Cited in DeMuth (1980b), p.30.

[10] This has been discussed in political science circles since at least Bernstein (1955). The imposition of RIA may also be seen as a mechanism of this kind; Spiller and Tiller (1997) set out a model that explains the introduction of RIA in the US by the Republican Congress as a way to control the pro-regulatory stance of the Clinton administration. Additional decision costs were imposed on agencies wishing to intervene, and additional grounds were created for courts to challenge such intervention.

[11] This cost measurement protocol, which is standardised to enhance its international applicability, is described in SCM Network (2005).

[12] The first study applying this approach to US federal data was Hahn and Hird (1991). More recently, the US Office of Management and Budget has been required to prepare a report each year by Section 624 of the Treasury and General Government Appropriations Act of 2001, termed the "Regulatory Right-to-Know Act," (Public Law 106-554, 31 U.S.C. 1105 note). The most recent final report is OIRA (2005).

The first option listed, regulatory impact assessment, involves putting in place formal, transparent processes by which regulators are required to assess the likely effects of proposed measures on social welfare. There are many models in use internationally, but core elements include:[13]

> "a) Appropriate problem definition and identification of policy objectives in such a way as to avoid ambiguities, vagueness and contradictions (with expected results expressed in quantitative, physical terms and an explicit hierarchy between objectives);
>
> b) Beginning of assessment when the choice is still open; consideration of multiple options;
>
> c) Information gathering - possibly through consultation - and data assessing, with an explicit choice of relevant criteria, procedures, and techniques for selecting a specific set of information;
>
> d) *Ex ante* impact assessment of each relevant option, through some explicit and consistently used method; description and most of the times quantification of effects; explicit selection of types of effects to be considered;
>
> e) RIA results expressed and publicized in a thorough and transparent way."

The other *ex ante* mechanisms listed vary in the extent of formality and transparency of application. Later in this section we will introduce the distinction between internally and externally focused evaluation models, and this distinction will also prove important in the selection of the right mechanisms for a particular jurisdiction.

---

[13] EU Presidencies (2004), p.11.

The *ex post* options involve applying additional oversight to regulators; for example, antitrust authorities increasingly see active scrutiny of regulatory restrictions as falling within their competence.

In principle, both *ex ante* and *ex post* mechanisms should improve the quality of regulation in the long-run. However, in practice there are likely to be important differences between their effects. *Ex ante* review may be essential for regulations that tend to be highly persistent once enacted. For example, measures that have the effect of transferring rents to particular groups may derive extra support if these "winners" are able to organise themselves in support of continuing the measures. Even after a negative *ex post* review, such measures may be hard to repeal.

Ex post review may have other advantages. First, it can be less informationally demanding, inasmuch as actual "before and after" outcomes or cross-jurisdiction benchmarks can be compared to the stated objectives of the measure and parameters that are particularly difficult to assess, such as the extent of compliance, can more readily be estimated.[14] Second, *ex post* review may be more amenable to being applied independent of the policymaking unit that proposed the measure. Such independent review may help address the incentive problems faced by agencies that both originate and assess measures.

[14] Ogus (1998), p.57, argues that cost-benefit analysis is less straightforward for regulation than for public programmes mainly because regulation is intended to change behaviour, requiring difficult predictions to be made about expected compliance.

2.4 Applying RIA efficiently in small regional jurisdictions

The design of suitable institutions to improve regulation in a small regional jurisdiction such as Ireland gives rise to many interesting economic questions. However, despite the extensive literature adverted to above concerning the shortcomings of regulation and potential responses to them, little consideration has been given to theoretical or empirical examination of the particular challenges in optimal design of regulation in small countries.

If we are societal welfare maximisers, we should prefer that the set of regulations applied be the one that provides the maximum net welfare benefit. It is widely understood that such an assessment should take all of the costs and benefits of each potential measure into account, including no-action options. However, it is worth emphasising that we must also take into account the effects on welfare of direct and indirect costs arising from the assessment process itself. The difficulty for a small regional economy is that "smallness" and "openness" may lead to higher unit costs of assessment.

To illustrate the effect of a jurisdiction's size, suppose that the cost of conducting a RIA is made up of a fixed component (e.g. for legislative drafting or setting up a survey) and a component that varies with the number of agents affected by the measure (e.g. more survey responses are needed because there are more types of agents). In a small jurisdiction, with few agents, the unit cost of scrutinising every measure will be higher than it would in a large jurisdiction, because the fixed costs will be spread over relatively few agents.

The unit cost of assessing regulatory impact might also be positively related to the trade-intensity, or openness, of an economy. In a highly open economy, it is likely that a greater proportion of possible regulatory measures will have a competitiveness impact that requires examination. If agents are prone to engage in international transactions and the

average intensity of such interactions is high, external conditions will tend to be of increased relevance in carrying out a RIA.

For a simple illustration of this effect, suppose that a proposed new fuel efficiency standard was under consideration for implementation in Ireland. If the standard was simultaneously being applied in the rest of the world, the main effects in Ireland might include changes in supply and demand conditions for goods in fuel-using sectors, plus general equilibrium effects. However, if the standard was intended for use in Ireland alone, with no simultaneous change in standards abroad, the expected impact in Ireland would include an additional set of effects driven by the increased relative price of Irish goods. Thus the effects of a single-country standard will tend be more complex to analyse than the effects of a global one; and analogously, the effects of a measure in a relatively open economy may be harder to measure than in a relatively closed one.

Given that deliberation is costly, and may be particularly costly in a small regional economy, various shortcuts are used to economise on it. In the remainder of this section we discuss four possible approaches:

- Analyse fewer measures, or analyse some measures less thoroughly;
- Vary the balance between internal and external focus;
- Improve the efficiency of the testing process by segmenting the analysis into general and specific tests depending upon the nature of expected impacts; and
- Use international data to economise on need for collection of unique data.

Below we discuss the implications of each of these potential responses.

Analyse fewer measures, or analyse each measure less thoroughly

Requiring some level of formal deliberation before regulations may be enacted inevitably gives rise to administrative costs. For example, resources required for regulatory impact assessment may include internal staff time, assistance from external experts, data collection processes and preparation of consultation responses by third parties. There are also likely to be social costs associated with delays to the subset of measures that were of good quality prior to assessment.

Such costs must be set against the welfare gains from improvements made to regulations. These benefits may include lower policy costs, avoidance of bad measures (or at least delay to them, if deliberation is imperfect), and welfare gains from adoption of more effective regulations.

Indeed, in an ideal system, additional deliberation would be undertaken only up to the point where its marginal cost equalled the marginal net social benefit of better regulation. This may mean that some measures should not be scrutinised at all. To see why, imagine that the cost of carrying out a RIA is invariant to the level of costs and benefits associated with a given proposed measure. If the population of regulatory measures that arise for consideration were to vary significantly in terms of costs and benefits, there might be cases where incurring the administrative cost of doing a RIA would not be justified given the (small) potential net impact of the intervention.

In practice, RIA systems incorporate features that allow implicit judgements to be made on how much deliberation is required for a given measure. First, the rules may define more than one level of general scrutiny, such as the initial, partial and full RIAs in the UK

system,[15] with a test in place to classify each measure's appropriate level. This test may include qualitative and quantitative components. Qualitative elements normally ask questions about the proposed measure designed to flag particularly important, politically sensitive or difficult measures. The quantitative component tends to take the form of a threshold on the expected gross administrative or social costs of the measure. Such tests are analogous to the market concentration thresholds commonly used in merger control regulation for exemption of "small" mergers from detailed review.

An example may be helpful. In the Irish RIA model, a "screening" RIA is required for a wide range of measures. To determine whether a more intensive "full" RIA is required, a mixed qualitative-quantitative test is applied:[16]

> "It is proposed that a Full RIA be conducted where any one of the following applies:
>
> (a) there will be significant negative impacts on national competitiveness;
>
> (b) there will be significant negative impacts on the socially excluded or vulnerable groups;
>
> (c) there will be significant negative impacts on the environment;
>
> (d) the proposals involve a significant policy change in an economic market;
>
> (e) the proposals will impinge disproportionately on the rights of citizens;
>
> (f) the proposals will impose a disproportionate compliance burden;
>
> (g) the costs to the Exchequer or third parties are significant, or are disproportionately borne by one group or sector. It is suggested that initial costs of

[15] Regulatory Impact Unit (2003).

[16] Department of the Taoiseach (2005b), pp.6-7.

> €10 million or cumulative costs of €50 million (to include both costs to the Exchequer and third parties) over ten years might be considered significant in this context…"

Setting thresholds for the level of scrutiny to be applied may be useful if there are reliable indicators for the likely impact of measures, or more particularly the likelihood that the expected welfare could be raised by further scrutiny. However, the essence of such a test is clarity and ease of application. If the wording of the test is ambiguous or employing it requires considerable specialised work, its effectiveness at reducing the administrative cost of RIA will be reduced.[17]

In addition to the *level* of scrutiny a measure should receive, the optimal *type* of scrutiny it should be given may also vary depending upon the nature of the measure. In particular, the effectiveness of some regulatory measures is heavily dependent upon their credibility, by which we mean governments' ability to credible pre-commit to a future course of action. For example, a regulated firm is likely to under-invest in efficiency improvements if it believes that government will subsequently use regulation to expropriate all (or too much) of its returns from such investments. If government can pre-commit to allow the firm a normal return on investment, higher social welfare may be achievable.

In cases where credibility is of particular importance, the choice of regulatory quality assurance mechanism must take this into account. Scrutiny mechanisms that tend to restrict regulatory independence, such as close control of independent regulators by policymakers, may reduce credibility and limit the scope for using regulation to place

[17] For example, in the Irish test cited above, the intended interpretation of terms such as "significant", "disproportionate" and "economic market" may need to be clarified in later guidance, as they are left undefined in the text.

strong incentives on firms.[18] In practice, this type of discrimination is often applied when institutions are put in place. Some sectors are deemed to require independent regulation, and independent regulators are often subject to different regulatory scrutiny mechanisms from those applied to central government departments.

Vary the balance between internal and external focus

The RIA models employed internationally differ along another important dimension: the extent to which evaluation of proposed policies is internal to the initiating department or agency or is validated by external review. We refer to these two types of models as *internally-focused* and *externally-focused*, but there is a continuum of options between the polar cases.[19] The choice of focus is likely to affect both the regulatory credibility and the practicability (including cost) of applying the RIA model.

Externally-focused methods for improving RIAs should in principle have an advantage in objectivity; external bodies may be more willing to identify shortcomings with a proposed measure than the originating agency, which may (in some sense) benefit from the policy's enactment. If the auditing body could obtain perfect information about the likely costs and benefits of the measure, we might prefer externally-focused evaluation to address this incentive effect. Note, however, that the auditors will have interests of their own that may be affected by the outcome of the evaluation process, so the incentives they face must be taken into account as well.

---

[18] The relationship between the institutional framework and credibility is discussed at length in Spiller (1996).

[19] Indeed, Radaelli (2005), p.934, points out that RIA models are sometimes applied in different ways within the *same* jurisdiction, depending upon the policy context.

In any event, the policy-making agency is likely to have an informational advantage over potential external auditors (or at least that the cost of transferring full information to auditors may be high). This means that the choice of evaluation strategy is not so obvious. Imposing tight external scrutiny will inevitably reduce the policymaker's autonomy and may provoke wasteful effort by the policymaker to exploit its informational advantage.

The appropriate balance between internally- and externally-focused scrutiny of regulation may thus depend on both institutional factors and informational ones. Radaelli (2005, pp.930-931) suggests that institutional factors help explain the diversity of RIA models internationally, and some of the factors he highlights are relevant to the feature of RIA we are discussing. Bureaucratic context (e.g. the prevailing degree of departmental autonomy) and whether a model is seen by policymakers as possessing legitimacy (e.g. not just a box-ticking exercise) are clearly important.

Informational factors are likely to vary across types of measure. For example, evaluating some measures requires sector- or topic- specific experts, whereas others are more accessible by generalists. This suggests that the optimal internal-external scrutiny boundary may vary by sector or subject of regulation, as well as by jurisdiction.

These potential trade-offs between internally- and externally focussed measures are recognised by policymakers; for example, Department of the Taoiseach (2005a) notes that the RIA model should have internal "legitimacy and support" but should also ensure that RIAs are "sufficiently rigorous."[20]

[20] Department of the Taoiseach (2005a), p.34.

Models closer to the externally-focused end of the spectrum can include external scrutiny at several points in the policy development process. For example, in the UK, several external requirements have their effects early in the policy-making process, including a formal cost-benefit test of whether a measure is justified,[21] a negative clearance filter for competition effects and a small firms impact test.[22] There is also formal *ex post* scrutiny of RIAs in the UK, via a detailed external review of a sample of RIAs each year by the National Audit Office.[23]

Internally-focussed models tend to be less prescriptive and to involve less formal external scrutiny, emphasising instead the goal of producing internal improvements in the policy development process. The new Irish model set out in Department of the Taoiseach (2005a, b) appears to be more internally-focused than the UK model. It places most of the responsibility for the RIA process on the department that is considering the policy measure. The RIA is clearly intended to be one input to policy among many, rather than imposing a formal test that must be satisfied before a measure may be implemented. Accordingly, the test of acceptability is more general; in principle, any objectives may be specified and any measure may be justified if it meets its stated objectives. There is (permissive rather than mandatory) scope for involving other departments or agencies in the preparation of a RIA. While the Taoiseach's Department and the Public Expenditure

[21] "All final RIAs should use the following wording in the Ministerial declaration: 'I have read the Regulatory Impact Assessment and I am satisfied that the benefits justify the costs.'," Regulatory Impact Unit (2003), p.37.

[22] Regulatory Impact Unit (2003).

[23] See for example National Audit Office (2005). Boyle (2005), suggests that the Office of the Comptroller and Auditor General could perform a similar role in Ireland (p.44).

Division of the Department of Finance have a loosely specified role in the RIA process, the model does not appear to include formal external review of RIAs.

Tension between internal goals (e.g. improving policymaking processes) and external goals (e.g. assuring the quality of final measures) is not unique to the regulatory sphere. McNamara and O'Hara (2005, p.267) describe how evaluation methods in the education sector tend to fall between "strict external control and inspection" and "self evaluation and internal regulation," and they argue that most systems tend to include elements of both methods, e.g. internal self-evaluation backed up by external monitoring.

Other options for finding the right balance between rigour and practicability include use of systematic peer review or formal transparency requirements.

Segment the RIA process between general and specific tests

Should the analysis of regulatory measures rely on a common methodology incorporating all possible effects, or do some types of regulatory effects require separate tests? A uniform test would have the merit of simplicity. However, particularly in a small jurisdiction, carrying out a full RIA of every regulatory measure would almost certainly involve excessive cost.

To mitigate this cost, we might exempt some measures from scrutiny or vary the rigour with which the test is applied, as discussed above. However, a further alternative is to identify a subset of effects for specific in-depth analysis, perhaps on foot of a preliminary screening test.

It might make sense to employ specific tests if some types of impact are expected to be pose a particularly serious problem in welfare terms **and** if there are few complementarities between measurement of these impacts and regulatory effects in general.

Part I of this thesis considers the appropriate design of one such specific test: a negative clearance test for competition impact.

Use international data to economise on need for collection of unique data

An option open to any jurisdiction, but of particular relevance to small regional ones, is to use international data on the impact of regulation as an input for RIAs. The added relevance of benchmarking approaches for such countries stems from the general pressure on them to economise on administrative inputs (discussed above), potential economies of scale in the collection of data and the advantages a small jurisdiction may have when looking for useful comparators.

Economies of scale in data collection arise in much the same way as do more general economies of regulatory administration. To the extent that there are fixed costs of conducting a survey or compiling its results, a large jurisdiction will face a lower unit cost of data gathering than a small one. In very small jurisdictions, there may also be added costs arising from a lack of specialist resources in the locality, requiring use of external specialists with resulting travel and other costs.

Small jurisdictions may also find it easier than large ones to find useful comparators. While there are a relatively small number of large countries with fully developed regulatory structures, small countries may be able to use regional data from within a bigger political unit as a comparator.

A related option for attaining greater economies of scale in administration is to identify policy measures used in other countries that can be copied in whole or in part, or to join supranational groupings and thereby "outsource" the relevant measures (which may include data collection).

Using international data to inform domestic regulatory choices has some disadvantages, but it can also have important benefits. Disadvantages include a potential lack of relevance, since markets and institutional contexts in two jurisdictions are never quite the same, and a risk that lessons will be misinterpreted by those with an imperfect understanding of conditions in the countries where the policy has been tried.[24]

However, the international data may have considerable advantages, if these problems can be addressed. If a regulatory measure has previously been employed in a sample of countries, careful analysis can provide actual evidence of its effects. In contrast, purely *ex ante* analysis can only ever provide a hypothetical view of a measure's impact (see for example the *ex ante* cost-benefit analyses of mobile number portability regulation discussed in Chapter 5). Some forms of *ex ante* analysis, such as bottom-up engineering models, are also likely to be considerably more costly to implement than international comparisons.

Part II of this thesis provides concrete examples of benefits and pitfalls associated with using international data to inform regulatory impact assessments.

2.5 Conclusions and summary

The term "regulation" takes a broad meaning in this thesis, encompassing economic and social measures intended to affect the behaviour of individuals and firms. Government interventions undertaken via taxation, public spending programmes and public investments may serve as substitutes for regulation, but evaluation of such measures is not discussed in detail in this thesis.

[24] Radaelli (2004) argues that trans-national comparisons of regulatory regimes should not involve importing decontextualised "best practice," but rather should draw lessons from international experience.

Partly due to wider institutional changes and partly to incentive effects, the use of statutory regulation has tended to grow relative to funded public interventions. While economists have put forward a variety of explanations for regulation, including mitigation of market failure and capture of rents by interest groups, there is a broad consensus that regulation is not always optimally applied in practice. In parallel with the growth of regulation and the development of theories about it, many countries have put in place measures designed to improve the quality of regulation.

A government wishing to maximise social welfare through its regulatory activities should try to maximise the net present value of regulatory costs and benefits, taking into account costs of administrative scrutiny, delays to desirable measures and interactions between the deliberative process and credibility. We have listed a range of *ex ante* and *ex post* mechanisms used or proposed to improve regulatory measures.

The optimum level and type of deliberation may vary by type of regulation (for example, it might differ between measures to control monopoly and measures to deal with externalities). However, optimal deliberation may also vary within such types due to the characteristics of the jurisdiction that is applying the regulation.

In particular, the size and "openness" of the competent administration may affect the nature and level of scrutiny that should be applied to regulatory measures. In a small regional jurisdiction, the cost per affected agent of efficient RIA may be higher for a given measure than it would be for a larger one.

We have discussed four possible responses to the particular challenges faced by small open jurisdictions.

One might analyse fewer measures, or analyse some measures less thoroughly. Some countries use threshold-based tests of significance to tailor the amount of deliberation a measure receives.

A second option is to vary the balance between internal and external focus, which involves choosing different points on the spectrum between rigour and ease of implementation.

A third option is to improve the efficiency of the testing process by segmenting the analysis into general and specific tests depending upon the nature of expected impact. One such test is discussed further in Part I of this thesis.

As a final option, to the extent that data collection and interpretation is also subject to economies of scale, small jurisdictions may wish to make more use of international data to economise on need for collection of unique data. Part II of the thesis contains some specific examples of how international data may inform regulatory parameters and of some of the difficulties that may arise when using international data in cases where its context is not fully understood.

PART I: INSTITUTIONAL MEASURES FOR ASSESSING REGULATORY IMPACT: TESTS FOR EFFECTS ON COMPETITION

The search for workable RIA shortcuts has led jurisdictions that employ this tool to adopt a wide variety of different approaches.[25] However, many RIA regimes share a common feature: specific tests are employed alongside a wider cost-benefit comparison in an attempt to capture particular types of impact that are thought to be of particular concern. Such specific tests may relate to effects on small businesses, the environment, charities and other voluntary organisations, churches, competitiveness, or a range of other areas.[26]

Our focus is this part of the thesis is on the design of one such specific test: the potential impact of regulatory measures on competition.[27] In Chapter 3, we examine why and how this sort of specific test might be applied. Taking the example of the UK competition filter (hereafter referred to as the "UK filter"), we identify some of the problems with an inappropriately constructed test and set out some ideas on how a better test could be constructed for a use in a small jurisdiction such as Ireland. Chapter 4 extends this analysis, showing that the UK filter is prone to "false negative" results for important classes of regulatory measures even when it is applied in the UK.

---

[25] For examples see Hopkins (1997) and EU Presidencies (2004).

[26] See EU Presidencies((2004), Part II (Country Experiences).

[27] We are not aware of a systematic treatment in the literature of the case for using specific tests rather than relying on a single all-encompassing test of regulatory impact, and it is not our intention to develop one here. However, we might suppose that jurisdictions use this approach for sources of costs or benefits that are considered likely to be significant (in at least some cases) and where the analytical tools used to assess effects are not expected to overlap significantly with those used for a broader review of regulatory impact. For other cases where a broader RIA review would be capable of incorporating specific costs and benefits, it might be better not to include a specific test, since this could lead to unnecessary duplication or imbalances in the treatment of different types of impact.

3 Testing which Proposed Regulations Need "Competition-proofing"[28]

There is plenty of evidence that regulatory measures can damage competition and reduce consumer welfare. We briefly review the main sources of concern in Section 3.1. In Ireland, which has recently introduced a system of regulatory impact assessment for use by central government departments,[29] there have been debates in recent years over the impact on competition of various regulatory and self-regulatory measures. Examples include the Groceries Order (e.g. banning below-invoice cost selling of groceries), aspects of pub licensing, planning restrictions on sizes of supermarkets, restrictions on the number of taxi licences, and Section 149 of the Consumer Credit Act, 1995 (which includes price controls on retail banks).

A thorough review of the competitive effects of a proposed measure may involve a significant amount of administrative effort. Specific data collection is likely to be required into issues such as the boundaries and structures of relevant markets, extent and nature of rivalry between firms, cost structures and so forth. These tasks require specialised expertise that may be scarce, particularly in smaller jurisdictions.[30]

---

[28] This chapter is substantially based upon Lyons, S., "Testing which Proposed Regulations Need 'Competition-proofing', *Journal of the Statistical and Social Inquiry Society of Ireland*, forthcoming. It was presented to a meeting of the Society on 26 January 2006. Small portions of that paper also appear in Chapter 1 and the introduction to Part 1.

[29] See Department of the Taoiseach (2005a) and (2005b).

[30] Moreover, if we assume that the administrative costs of subjecting a measure to a given level of scrutiny are broadly fixed, while the benefits (and some other costs) tend to depend upon the number of persons or firms affected, the net cost per taxpayer for a scrutinising a given measure is likely to be higher in a small country than in a large one.

Given that many measures are purely administrative in nature or otherwise unlikely to have any effect on competition,[31] it would be wasteful to require a detailed competition review for all proposed measures. A negative clearance test may be more efficient – preferably one that is easy to apply but discriminates reliably between measures that need to be examined and those that do not.

The UK "competition filter" is probably the best documented test of this kind, so we take it as our starting point. However, we find that the UK filter has important shortcomings – in essence, there are classes of potentially damaging measures that will not trigger it. Our focus is on finding an appropriate test for a small country such as Ireland, but our results also point to weaknesses in UK regulatory impact assessment policy.

Identifying the problems with the UK filter helps us to outline an alternative type of filter that should better discriminate between safe and potentially damaging measures, while also being administratively practicable to apply.

3.1 Why test the effects of regulatory measures on competition?

We assume throughout this chapter that the appropriate goal of a policymaker, whether a legislator or civil servant, is to improve social welfare. This assumption places the analysis within the "public interest" approach to government intervention, in a normative sense.[32] However, we recognise that this motive for action may not prevail in all cases: there is potential for capture of agency decisions by interest groups. This means that

[31] A review of RIA in Canada reported that "About 30 percent of regulations are administrative in nature and have almost no economic impact." Treasury Board of Canada (1997), p.5.

[32] We have already – in Chapter 2 – encountered the positive distinction between the "public interest" (or NPT) and "capture" approaches to explaining public intervention in markets. For a summary of the distinctions between these schools see Laffont and Tirole (1993), Chapter 11.

institutions constructed to maximise the public interest need to account for, and be robust to, the risk of capture.

To see how regulatory measures can reduce welfare, and thus identify the sorts of measures welfare-maximising governments should prefer not to take, we first consider the reasons that competition-reducing regulations are adopted and the ways in which they may reduce social welfare. The framework outlined in this section will be useful later in the chapter for identifying the characteristics of measures that are likely to cause significant harm to competition.

Later in this chapter we look at some of the administrative measures that governments use to reduce the likelihood that damaging regulations are adopted.

Reasons for adopting measures that will harm competition

We can distinguish between two sorts of reasons that a policymaker might have for adopting a regulatory measure while knowing that it will have a negative impact on competition. The first sort, which we describe as **trade-offs**, is consistent with a public interest approach, in that it includes justifications that a benevolent social-welfare maximising policymaker could accept. Trade-offs are justifications that involve taking actions that will improve net social welfare, but in so doing will also cause some unavoidable harm to competition in one or more markets. Policymakers and interest groups often invoke trade-offs to justify proposed measures, and sometimes these arguments are legitimate. Examples abound; perhaps the most obvious ones are environmental or safety measures that improve welfare by correcting for externalities, but also have the effect of imposing significant sunk costs on a sector. This may in turn affect the conditions of market entry and hence reduce competition. But if the welfare gains

from the measure are sufficiently strong and there is no lower-cost alternative, the policy may well be justified.

The main risk to welfare from measures involving legitimate trade-offs is that the relevant costs and benefits will be miscalculated. If the error is significant, harm to social welfare could be correspondingly large. This issue arises for all types of costs and benefits, not just those associated with competition, but competition analysis could be used to reduce the level of likelihood of error for some measures.

The second potential class of reasons we identify for adopting measures that are expected to harm competition is that involving **rent capture**. As noted in Chapter 2, an extensive literature following Stigler (1971) sets out an economic theory of regulation, whereby well-organised minorities obtain rents by capturing political and regulatory institutions. Of relevance to our analysis is Stigler's original classification of the types of policies through which capture may be attempted.[33] Two of the four channels he identifies are of lesser relevance to our analysis. The first of these is direct government subsidies, and the second includes policies that support goods complementary to those of the lobbyists (or hinder substitute goods). However, the other two types of policies Stigler identifies involve regulatory measures.

The first, and probably most important of these two types, is creation of entry barriers using mechanisms such as licensing or standards. Such policies are particularly desirable for the rent-seeker because they can ensure that any rents captured will be retained by existing market participants.[34] The second type set of relevant policies involves the

[33] Stigler (1971), pp. 4-6.

[34] Laffont and Tirole (1993), p.504.

imposition of price controls that serve to weaken price competition. Such policies are particularly relevant in markets that are already subject to barriers to entry, either due to other administrative measures or because of structural characteristics.

Capture may lead to larger and more persistent harm to welfare than mere errors in estimating the value of legitimate trade-offs. The damage to welfare from this source is likely to include the direct distortion arising from the creation of rents, *plus* deadweight losses associated with the process of capturing and retaining them. In addition to static losses, there may be a dynamic effect: potential for capture of rents is likely to elicit continuing effort to obtain them.[35] Thus if we find measures with characteristics that may indicate a risk of rent capture, we should require strong justifications before allowing them to be imposed.

Ways in which measures may damage competition

Regardless of the reasons for their adoption, measures can limit the scope for competition either by adversely affecting the incentives for entry, expansion or exit of firms, or by weakening the rivalry of firms already in the market. Below we briefly review each of these classes of impact.

Limiting scope for entry, expansion or exit from markets

The most obvious way regulation can limit entry is by imposing explicit rules that govern the number or identity of market participants. The statutory public utility monopoly is a familiar, if increasingly rare, example. Even where there is no public monopoly, the state

[35] There is a useful survey of rent-seeking and measures used to combat it in Milgrom and Roberts (1992), pp.270-284.

may impose quantitative limits on the number of firms allowed to serve the market, often through some form of licensing arrangement.

Regulation can also have an indirect effect on terms of entry, expansion or exit. For example, government actions that increase the cost of doing business, particularly where higher costs are imposed on actual or potential entrants than on incumbents, can reduce the number or effectiveness of entrants.[36]

Weakening rivalry

Competition may be more or less fierce for a given number of competitors in the market, and regulation may affect the level of rivalry just as it can limit entry.

Here too, direct and indirect effects are possible. Direct limits to rivalry can take the form of limits on firms' pricing behaviour, advertising, quantities offered, variety or product characteristics. Indirect effects may arise if regulation serves to reduce firms' uncertainty about competitors' likely strategic choices, making it easier for market participants to actively or passively coordinate behaviour.

Administrative measures to reduce damaging regulation

Without knowing more about a specific regulatory measure, we cannot be sure if it will increase consumer welfare or reduce it. Even restricting ourselves to the subset of measures that have some negative impact on competition, a given measure may:

- Increase welfare, if the net benefits arising from the measure are positive, taking into account the damage to competition; or

[36] Asymmetries of this kind may be explicit, as in the case of "grandfathering" provisions, or implicit, such as in cases where the cost structures of entrants and incumbents are materially different, leading to asymmetry in the incidence of a regulatory measure.

- Decrease welfare, if net benefits are not sufficiently large (e.g. the main effect of the measure is to appropriate rents for a particular group).

To help increase the likelihood that adopted regulations are welfare-increasing, some jurisdictions require policymakers to carry out additional analysis on the types of measures that are deemed to give rise to high risk of damage to competition.[37] Such tests are analogous to merger control, whereby a class of transactions are identified that give rise to particular risks of harm to consumers, and the regulator is required to examine them using suitable analytical tools.[38] Because mergers are difficult to reverse after completion, reliance on *ex post* controls would not be sufficient to avoid persistent harm: *ex ante* examinations are required.[39]

Adoption of a regulatory measure might seem a more reversible process than a merger, but the very measures that cause the greatest harm are likely to have the strongest tendency towards persistence: those involving capture of rents by a small, well organised group at the expense of the wider society. Even in the absence of such political resistance to reversal, reversing a regulatory measure may involve significant administrative cost and delay. Indeed, there may even be cases where reversing the outcome of a damaging regulation could be more costly than leaving the outcome in place. An example of this might be imposition of a sub-optimal technological standard; getting back to the original first-best position could involve prohibitive adjustment costs.

---

[37] In the case of primary legislation, legislators are essentially constraining themselves by adopting RIA. With reference to secondary legislation and actions by regulatory agencies, there is an extensive literature viewing the adoption of administrative procedures such as these as a way for legislators to manage principle-agent problems in the regulatory process. See for example McCubbins *et al*. (1987).

[38] The EC Merger Regulations, set out in European Council (2004), provide a recent European example.

[39] Recitals 5 and 7 in the Merger Regulations, European Council (2004), make it clear that *ex post* competition law measures were considered insufficient to prevent potential harm from some concentrations

Of course, measures invariably have effects that go beyond an impact on competition. Policymakers normally embed competition analysis in a wider process of regulatory impact assessment. While it would be wrong to ignore the impact of a new measure on competition, it would be equally foolish to ignore the other sources of costs and benefits associated with the measure.

Indeed, since a large proportion of legislative measures are purely administrative in nature or for other reasons have no effect on competition, it would seem excessive to prescribe that all proposed legislative and regulatory measures be competition-proofed. One solution to this problem, again analogous to merger control, is to use a negative clearance test to quickly dispose of measures that are not going to raise any significant concerns.

Probably the best documented negative clearance test for the competition effects of regulations is the UK competition filter developed by the Office of Fair Trading. In the next section, we describe this mechanism and consider how well it meets its stated objectives. While this analysis has obvious relevance to the UK, where the system is used, we are particularly interested in how well a system of this kind would perform in a smaller jurisdiction such as Ireland.

3.2 The UK competition filter

The UK test for whether a regulatory measure undergoing RIA requires extra scrutiny for competition effects is called the "competition filter." While it is widely used for RIA studies in Britain and is accompanied by extensive documentation, we find that it has significant theoretical and empirical/applied shortcomings.

In this section, we describe the UK filter briefly, and then apply it to four measures that are – or were recently – on the statute books in Ireland. These measures were chosen because each has at some point been alleged to damage competition significantly and

undergone some form of impact examination. We find that at least three, and possibly all four, of the measures would **not** have been required to undergo competition scrutiny if the UK filter were applied to them.

Drawing on these examples and our earlier discussion of ways regulation can damage competition, we identify major problems with the UK filter. In the next section, we apply these lessons and try to outline a more appropriate competition filter for use in RIA exercises.

Description of the UK filter

The UK filter requires a policymaker to consider nine yes/no questions. If the answer to five or more of these questions is yes, a review of the competition effects of the proposed measure is deemed necessary. The rationale for the UK filter is set out in OFT (2002), together with detailed guidance for policymakers on how it should be applied and how a detailed review should be conducted where it proves necessary. In this section, we first list the questions in the filter and then summarise the guidance provided on each question in OFT (2002).

The nine questions are listed in Table 1 below.

Table 1: UK competition filter questions
1. In the market(s) affected by the new regulation, does any firm have more than 10 per cent market share?
2. In the market(s) affected by the new regulation, does any firm have more than 20 per cent market share?
3. In the market(s) affected by the new regulation, do the largest three firms together have at least 50 per cent market share?
4. Would the costs of the regulation affect some firms substantially more than others?
5. Is the regulation likely to affect the market structure, changing the number or size of firms?
6. Would the regulation lead to higher set-up costs for new or potential firms compared with the costs for existing firms?
7. Would the regulation lead to higher ongoing costs for new or potential firms compared with the costs for existing firms?
8. Is the market characterised by rapid technological change?
9. Would the regulation restrict the ability of firms to choose the price, quality, range or location of their products?

Source: OFT (2002).

The filter is intended to "quickly signal those proposals that are most at risk of impacting materially on the competitive process"[40] and it is to be carried out early in the process of developing policy. It is clear from the guidance that the filter is to be applied to economic markets, and a simple description is given of how such a market should be defined: "a market includes the firms that compete against one another to sell the same or similar products or services."[41] The guidance stipulates that the filter should be applied to all affected markets, or in the case of regulations with broad sectoral impact, at least those

[40] OFT (2002), para. 4.3.

[41] *Ibid*, para. 4.7.

markets likely to be affected most intensely. However, it is clear from the text that users of the filter are not expected to carry out full-scale reviews of market boundaries.[42]

The OFT briefly describes the logic behind each of the nine questions. The questions can be grouped into two sets: four questions about market characteristics (namely Questions 1, 2, 3 and 8) and five that focus on the nature of the regulatory measures under consideration (namely Questions 4, 5, 6, 7 and 9). To see how the filter is constructed, we outline the rationale given for the questions, taking the market-related ones first.

Three of the nine questions in the filter relate to the *ex ante* level of market concentration. Questions 1 and 2 ask whether the market share of the largest firm (in terms of UK sales) is at least 10 or 20 per cent. The guidance states that market power is more likely to be present where one or two firms have a significant share of the market and that higher market share would be "more of a concern."

The third question is closely related to the first two, asking whether the largest three firms have at least a 50% share. This is taken to imply a greater risk both of single firm market power and potential scope for collusion – "cases where a few large firms may be able to act together."[43]

The eighth question also relates to a characteristic of the market, asking whether it is subject to "rapid technological change." Such markets are singled out due to the "risk that regulation may restrict innovation in such markets."[44]

---

[42] As discussed later in the chapter, omitting a requirement for full-scale market reviews is probably a practical necessity.

[43] *Ibid*, para. 4.19.

[44] *Ibid*, para. 4.30.

We now turn to the remaining five questions, which focus on the nature of the measure.

Potential for asymmetric incidence of the measure is the focus of Question 4, which asks if "the costs of the regulation [would] affect some firms substantially more than others."[45] The OFT explains that the firms considered must be ones competing with one another, and three possible sources of asymmetric costs are suggested: small firms affected differently from large ones, firms disadvantaged because of "the resources they use" and firms in some locations affected differently from those in other places. It appears from the text that these must be existing firms, as opposed to potential entrants. In judging whether the asymmetry is "substantial," a policymaker should consider whether firms disadvantaged by the measure could nevertheless remain in business, or whether they might exit the market.

Question 5 asks whether the measure is likely to "affect the market structure, changing the number or size of firms."[46] From the brief guidance given, it seems that this relates to the likelihood of market consolidation or exit by firms as a result of the measure. There is no indication that the answer should reflect any potential for deterrence of future entry, which, as we shall see later, is an important caveat.

The sixth and seventh questions ask whether there would be higher set-up or ongoing costs for actual or potential entrants than for existing firms. Specific reference is made to the possibility that licensing or restrictions on location might be applied, inhibiting entry, or that treatment of incumbents and entrants might be asymmetrical in some way.

---

[45] *Ibid*, para. 4.20.

[46] *Ibid*, para. 4.23.

Finally, question 9 asks whether the measure includes other forms of economic regulation: restrictions on prices, quality, range or location of products. Examples given include product standards, price controls and permitted quantities of inputs. The guidance comments that "all will have the effect of removing one way in which firms can compete, and therefore represent a distortion to competition."[47]

In summary, the questions regarding the nature of the measure try to capture, in a cumulative way, the possibility that the measure will change market structure, raise various barriers to entry and impose restrictions on firms' competitive behaviour.

Four examples of measures that the UK filter might exempt from scrutiny on competition grounds

In this section, we look at four actual regulatory measures that are currently in place in Ireland, or were in place in the recent past, and which are or were alleged to have potentially significant effects on competition. Our objective is not to determine if these measures are justified or not, but to use them as stylised examples for illustrating how particular types of policies would fare under the UK filter. The underlying assumption is that the filter should flag possible concerns if such measures such as these were proposed.

In three of these cases, the filter suggests that the relevant measure *does not* require competition scrutiny. The result in the fourth case is ambiguous, but a case could be made for passing it.

This conflict between our prior expectations and the results of applying the UK filter helps to illustrate the shortcomings of the UK filter, and by extension, to highlight some of the

[47] *Ibid*, para. 4.32.

features required for a better negative clearance test for the effects of regulation on competition.

Example 1 – Limitation on the number of licences issued in the Dublin taxi market

Between 1978 and 2000, restrictions were imposed on the number of taxis licensed to operate in Dublin and other parts of Ireland.[48] Other concurrent measures imposed price restrictions, qualifications for drivers and vehicle standards. As the relevant markets are likely to be local or regional in scope, we focus on the Dublin market.

Suppose the quantitative restrictions had been renewed rather than lifted in 2000, but that they were also subjected to a regulatory impact assessment. Would these measures have been captured by the UK filter? As an aside, we note that although these restrictions have been lifted, taxi regulation remains a live issue: new regulatory structures are currently being developed for the sector.[49]

Our prior expectation is that the measure should trigger the filter. Research at the time the restriction was in place indicated that it imposed a high welfare cost by restricting the supply of taxis. One estimate put the monopoly rent associated with the restrictions at over €15 million per annum.[50] Long queues were commonplace.[51]

---

[48] This restriction was imposed via secondary legislation, Statutory Instrument No. 292 of 1978, cited in Fingleton *et al.* (1998), p.4. It was removed on 21 November 2000, following a High Court decision: OECD (2001a), p.28.

[49] The legislative framework is set out in the Taxi Regulation Act, 2003.

[50] Fingleton *et al*. (1998), p.11.

[51] See, for example, Kaminski (1999).

However, application of the UK filter (see Table 2 below) suggests that this measure would not give rise to competition concerns and that its competition impact should therefore have been exempted from a detailed examination.

Since five positive responses are required to signal a need for reviewing the competition impact, and application of the filter yields at most two, this measure would apparently be waved through.

Table 2: Application of UK filter to Example 1: ***Limitation on the number of licences issued in the Dublin taxi market***		
Question(s)	**Likely Answer**	**Reasons**
1-3	No	The market was not considered to be concentrated; there were 2,374 taxi licences in place in 1998, and most were owned by owner-drivers.[52] Some individuals and firms held sets of licences, but according to Fingleton, Evans and Hogan (1998), the share of licences held this way was "probably considerably less than 25 per cent." Hackneys, numbering about 3,000 in 1998, also served a segment of the market.
4	No	There is no reason to think that the costs of regulation affected existing firms in an asymmetrical way.
5	Yes	The measure had direct effects on the market structure, limiting the number of market participants.[53]
6	Yes	Potential entrants would have to purchase an existing licence from an incumbent to be able to enter.
7	No	There were no apparent differences in ongoing costs for different types of competitor.
8	No	This was a mature market without rapid technological change.
9	No	The regulation did not directly restrict price, quality, range or location of services. However, other regulatory measures applied to taxis did affect such service characteristics.

Observations on this example

The measure passes mainly because the relevant market is not found to be concentrated. The cost asymmetry imposed by the regulation is highly focused, and the filter misses the point that barriers to entry can be narrow in focus but powerful in effect. Indeed, the filter effectively regards measures that affect both start-up and on-going costs asymmetrically

[52] Fingleton *et al*. (1998), p.6-7.

[53] This response assumes that the measure is being compared to a no-restriction counterfactual. However, if the measure were to be considered for retention, a case might be made for a 'no' answer to Question 5 on the basis that retaining the restriction involves no change in the already-restricted market structure.

for incumbents and entrants as twice as deserving of attention as measures that affect only one type of cost, regardless of the intensity of impact or extent of asymmetry.

Entry regulations affecting non-concentrated markets are not uncommon, and it seems likely that most such measures would pass the UK filter. For example, many such measures have been identified in the Irish Competition Authority's ongoing review of selected professions.[54]

Example 2: Restriction on the eligibility of overseas-trained pharmacists to open a pharmacy in Ireland.

Under Irish law (and that of several other EU countries), pharmacists trained outside the State are not permitted to manage or supervise a pharmacy unless it has been in operation for more than three years.[55] The OECD has recommended that the restriction be lifted to facilitate competition,[56] and Purcell (2004) cites it as one of the two most important barriers to entry into Ireland's retail pharmacy market.[57] Whether or not there is an adequate "trade-off" justification for such restriction, if a measure like this were subject to a regulatory impact assessment, competition effects should be examined.

Again, we apply the UK filter to the measure (see Table 3 below).

---

[54] See http://www.tca.ie/professions.html.

[55] Purcell (2004), pp.40-41. The Irish legislation is the European Communities (Recognition of Qualifications in Pharmacy) Regulations, 1987, 1991 and 1994, transposing EC Directive 85/433/EEC. In 2006 the Irish restrictions survived a court challenge that was appealed as far as the European Court of Justice (case C-221/05); see Smyth (2006).

[56] OECD (2001a), pp.88-89.

[57] Purcell (2004), p.*xii*.

Table 3: Application of UK filter to Example 2:
Restriction on the eligibility of overseas-trained pharmacists to open a pharmacy in Ireland

Question(s)	Likely Answer	Reasons
1-3	No	The (national) market is not considered concentrated; the largest firm has a 4% share.[58] Of course, if markets were considered to be local, some might be found to be more highly concentrated.
4	No	There is no reason to think that the costs of regulation would affect existing firms in an asymmetrical way.
5	Yes	Possible effects on market structure would be a concern.[59]
6	Yes	There would be higher set-up costs for certain types of potential firm (i.e. those that would have been run by an overseas-trained pharmacist). Indeed, the set-up costs for these firms could be said to be infinite.
7	No	We can probably conclude that there are no extra ongoing costs for new or potential firms, given that firms affected by the regulation will simply not be permitted to enter.
8	No	This appears to be a relatively mature market without rapid technological change. Of course, the rate of change may have been affected by the presence of the restrictions, so the counterfactual that is chosen may affect this question.
9	No	The regulation does not directly restrict price, quality, range or location of services.

Application of the UK filter yields two 'yes' answers (or at most three, if we took Question 7 to merit a 'yes'). The measure would clearly pass the UK filter.

[58] *Ibid.*

[59] As with Question 5 in Example 1, the answer to this question might be 'no', depending upon the counterfactual being applied.

Observations on this example

Once again, the main reason for the measure passing the filter is that the relevant market is not deemed to be concentrated. The focused nature of the barrier to entry imposed by the measure, together with maturity of the market, contribute as well.

In this case, it is also interesting to note that the lack of price controls or other remedies in the measure helps obtain a no-scrutiny result. At first glance, this makes sense: we have noted earlier that price controls may have a significant effect on competition, even potentially capturing rents. However, treating entry and price control features in a parallel way leads to the perverse result that a policy that limits entry but not prices is treated as less likely to damage welfare than one in which both entry and price are controlled.[60] Yet restricting entry in a market with unconstrained prices can amount to a rent-seekers' charter.

If entry regulation proves to be necessary, consumers may be better off if prices and quality are regulated as well. Thus regulatory measures that control entry alone seem at least as deserving of examination as those that control both entry and prices.

Example 3 –Ban on selling groceries below invoice cost

Our third example involves a very different sort of measure and market, yet it too passes the filter despite having repeatedly attracted scrutiny in Ireland on grounds of possible effects on competition.[61]

[60] I am indebted to Greg Swinand for discussions on this issue.

[61] Several of these previous reviews are cited in CMRG (1999), p.7.

This measure is the Restrictive Practices (Groceries) Order, 1987 (hereafter described as "the Order").[62] In this examination we focus on one element of the Order:[63] its prohibition on selling various grocery products below invoice cost:

> *Under Article 11 a retailer is prohibited from selling grocery goods (except in very limited circumstances in the case of goods whose date of minimum durability has expired) which is less than the net invoice price of the goods (including value added tax) or where charges in respect of carriage, insurance and other costs not included in the relevant invoice have to be paid by the retailer to the supplier to obtain delivery of the goods to his premises. The amount obtained when such charges or costs are added to the net invoice price of the goods (including value added tax).*[64]

Critics of the Groceries Order argue that it damages competition by requiring grocery retailers to charge prices higher than the actual price of goods (once off-invoice discounts are taken into account). In effect, they say, it amounts to a form of resale price maintenance.[65] Although supporters of the measure believe it would bring important benefits, it is clear that competition effects of a proposed measure similar to the Groceries Order would merit examination.

The filter is applied below, based on information about market conditions reported in CMRG (1999) during a past review of the measure.

---

[62] The Order remained in force until revoked by the Competition (Amendment) Act, 2006.

[63] The Order also includes restrictions on a range of other practices, e.g. resale price maintenance, "hello money," refusal to supply, advertising and various types of agreements. Some of these restrictions were retained in the 2006 Act cited above, but none of them appears to change the results of the filter.

[64] CMRG (1999), pp.10-11.

[65] For three examples, see CMRG (1999), p.27-31.

Table 4: Application of UK filter to Example 3: ***Ban on selling groceries below invoice cost***		
Question(s)	**Likely Answer**	**Reasons**
1-3	Yes	If we take the relevant geographical market to be national and the relevant product market to be groceries, one firm had 21% of the market in 1998 and the top three firms had 60% share.[66]
4	No	There is no reason to think that the costs of regulation would affect existing firms in an asymmetrical way.
5	No	The measure should not cause changes in market structure; indeed, its advocates argue it is needed to prevent abuse of market power that would lead to further consolidation.
6	No	No obvious effect on set-up costs.
7	No	No obvious effects on ongoing costs.
8	No	There is some technological change in the grocery retailing market, but we suspect that most generalist observers would not characterise it as rapid.
9	Yes	The regulation restricts pricing behaviour.

In this case, the three concentration-related questions would all be likely to yield "Yes" results. The measure also involves a clear restriction on pricing, which falls under Question 9.

However, the measure does not impose asymmetrical costs or bear differently on incumbents than entrants. The rapidity of technological change in grocery retailing is debatable, but it seems unlikely that most observers would provide a 'yes' answer to this question since the sector is not *prima facie* technology-intensive. Moreover, in circumstances where the restrictions were in place and under review, it is possible that the

[66] CMRG (1999), p.55.

rate of technological change might have been reduced by the presence of the restrictions. This could lead to ambiguity with the counterfactual that should be applied similar to the issue we encountered for Question 5 in Example 1 and Questions 5 and 8 in Example 2 .

In any event, the net result falls short of the five "Yes" answers needed to trigger scrutiny if a policymaker were to propose a measure such as this.

Observations on this example

Even in a concentrated market, it seems that a measure that is focused on one aspect of firms' behaviour, no matter how intense its effects, may pass through the filter. Indeed, it seems from the examples we have tested that the only type of intervention likely to be captured by the filter is one for which the relevant market is concentrated (*ex ante*) ***and*** the measure acts through a range of economic channels (e.g. affecting both entry and pricing or affecting entry in a range of ways).

One suspects this result is not unique to the Groceries Order. If we were to apply the filter to other narrow but potentially harmful measures affecting concentrated markets, it seems likely that the result would be the same.

Example 4 – Price controls on retail banks

Our final example has some similarities to the third one, in the sense that the relevant market would probably be found to be concentrated under the UK filter. However, the answers to the other questions are more ambiguous in this case than in our other examples, and the nature of the ambiguity highlights the scope for exercise of discretion when applying the UK filter.

Section 149 of Ireland's Consumer Credit Act, 1995, *inter alia* imposes a price control regime on a range of products offered by retail banks (or credit institutions, to be more precise). If a regulated firm wishes to introduce a new service or increase the charge on an

existing one, the proposed charges must be notified to the financial services regulator in advance, together with an administrative fee and supporting information. The regulator then determines whether the charge is acceptable.

A study carried out for the Competition Authority argues that regulation of bank fees under Section 149 "increases costs, creates risk, and limits commercial freedom for banks. All of these effects serve to discourage entry by new providers and innovation by existing providers, and consequently they create harm to consumers."[67] We would certainly expect that any effective competition filter would pick up a detailed price control measure such as this.

As in the previous examples, we apply the filter below.

[67] LECG (2004), pp.49-50.

Table 5: Application of UK filter to Example 4:
Price controls on retail banking services

Question(s)	Likely Answer	Reasons
1-3	Yes	Data published in LECG (2004) suggests that if the relevant market were taken to encompass personal current accounts in Ireland, concentration would exceed all three thresholds in the UK filter.[68]
4	No/Yes?	The measure confers discretion on the regulator as to the administrative charges that may be applied.[69] The answer to this question is therefore ambiguous; costs might or might not be asymmetrical depending upon how the regulator applied the measure.
5	No	The measure should not cause changes in market structure; like other measures discussed in this chapter, it is more likely to hinder structural change than induce it.
6	No/Yes?	Set-up costs are open to regulator discretion (see Q4 above). All new services must be submitted for price approval, but application charges are discretionary.
7	No/Yes?	Ongoing costs are open to regulator discretion (see Q4 above). All price increases or new product introductions must be submitted for approval, but application charges are discretionary.
8	No/Yes?	While there is a degree of technical change in retail banking, it is less obviously innovative than technology-driven sectors like computing and biotechnology. The answer to this question would depend upon the views of the person completing the filter.
9	Yes	The regulation restricts pricing behaviour.

[68] LECG (2004), Table 2, p.27.

[69] For each price application, the regulator may charge "such fee as the Director may decide as respects each notification, being not more than a fee of £25,000 [€31,750] or such other amount as may stand specified in regulations."

It seems likely that this market would be deemed concentrated under questions 1-3 of the filter, and being a price control measure it obviously merits a "yes" for question 9. However, the answers to most of the other questions are ambiguous, and in all cases could easily be deemed to be 'no'. If all of these were answered in the negative the measure would pass the filter.

It would not be difficult to construct an argument for a "no" for each of the four ambiguous questions. In considering questions 4, 6 and 7, note that the regulator has discretion as to the administrative costs it may impose. The legislation states:

> The fee referred to in subsection (3) may be waived or reduced by the Director where the payment of the fee would, in the opinion of the Director, be unfair to the credit institution having regard to—
>
> (a) the impact of any increase in or imposition of charges on customers,
>
> (b) the number of customers affected by any increase in or imposition of charges,
>
> (c) the additional income likely to accrue from any increase in or imposition of charges, and
>
> (d) any other criteria that he deems appropriate.[70]

A policymaker keen on the proposal might assume that the regulator would exercise this discretion in a competitively neutral way, and that other costs arising from the measure (e.g. preparation of pricing applications) would have broadly symmetrical effects across regulated firms. Alternatively, a more sceptical application of the test could yield a "no" for each of these questions, for example by assuming that administrative charges would be

[70] Section 149 (4) of the Consumer Credit Act, 1995.

applied uniformly and that a new entrant would have a much higher frequency of new service applications and price change applications while developing its set of offerings in unfamiliar market conditions.

Finding a way to answer "no" to question 8 is more straightforward still: there is no absolute metric in the test for whether or not a market is subject to "rapid technological change".

Observations on this example

This ambiguous result may be as problematic as a straightforward passing result for two reasons. First, the policymaker applying the filter has the (perhaps) unenviable task of anticipating the effects of future exercise of discretion by an agency. This may require specialised economic knowledge. Perhaps this problem could be mitigated by providing more detailed guidance (e.g. basing the test on a worst-case scenario about discretionary actions), but this would add further complexity to the filter.

The second problem is that ambiguity heightens the risk of capture by interest groups, as discussed in Section 3.1 above. Capture might take the form of a "favourable" application of discretion in carrying out the test. Furthermore, the wide discretion implied by these questions might provide an added incentive for proponents to focus rent-seeking efforts on the policymaker.

Summary of examples

In the table below, we summarise the examples discussed in this section.

Table 6: Summary of examples (✓ = yes, - = no)				
Question	**Example 1** *Restriction on number of taxi licences*	**Example 2** *Rule on foreign-trained pharmacists managing pharmacies*	**Example 3** *Ban on below-invoice cost selling of groceries*	**Example 4** *Price controls on retail banking services*
1	-	-	✓	✓
2	-	-	✓	✓
3	-	-	✓	✓
4	-	-	-	(- / ✓)
5	✓	✓	-	-
6	✓	✓	-	(- / ✓)
7	-	(- / ✓)	-	(- / ✓)
8	-	-	-	(- / ✓)
9	-	-	✓	✓
Overall result	**Pass** (Yes = 2)	**Pass** (Yes = 2-3)	**Pass** (Yes = 4)	**Ambiguous** (Yes = 4-7)

Source: analysis by the author.

We can distinguish between measures likely to pass mainly because the market will not be deemed to be concentrated (e.g. examples 1 and 2), and measures that pass mainly because the nature of the economic effect is (or can be presented as) narrow, despite a significant level of market concentration (e.g. examples 3 and 4).

Conceptual problems with the UK competition filter

The UK filter does not appear to be well suited to its intended function. In particular, it seems likely to allow through some measures with high risk of capture and damage to welfare. Both the structure and detailed content of the filter seem to contribute to this problem; we take each in turn.

Note first that mechanism has a *parallel* structure, in the sense that it attempts to aggregate a multi-dimensional set of market and regulation attributes in a single calculation. Applying criteria in this parallel way implies that there are important trade-offs between them, and that they can be aggregated meaningfully. Moreover, the filter combines questions about the nature of the market with questions about the nature of the measure under review. This means the implied trade-offs and weightings are very complex and hard to fathom. For example, does less concentration really equate to the absence of price regulation in driving the likelihood that a measure will harm competition?

A key problem arising from the parallel structure of the filter is that no single feature of a measure or market, no matter how significant, can lead to a positive (or negative) result. An alternative *serial* (i.e. step-by-step) approach, which takes a involves a series of binary decisions, would probably be better at picking up black and white cases, as opposed to those involving shades of grey.

Parallel structure may also lead the UK filter to give perverse incentives to policymakers, for example by encouraging them to introduce measures on a piecemeal basis. If policies with narrow economic effects are automatically deemed less likely to require a detailed competition assessment, a policymaker wishing to get a broad measure through while minimising the administrative burden might proceed with only part of the package, and not necessarily the best part. An example mentioned earlier is the apparent incentive for

introduction of entry restrictions without matching price controls – precisely the outcome that would favour capture of rents.

Turning to the content of the UK filter, a key problem arises from its focus on concentrated markets. First, high market concentration is not a necessary or sufficient condition for a regulatory measure to damage competition. Second, if there are cases where high concentration increases the likelihood that a measure will harm competition, it must be the level of concentration *after* regulation is imposed that matter, not concentration *ex ante*. Of course, *ex ante* concentration is examined in the application of some specific regulatory measures, such as horizontal merger control. But its importance in this context is as an indicator of likely post-merger concentration.

Markets that are concentrated before they are regulated are hardly the only ones, and perhaps not even the main ones, in which regulation may harm competition. To take a simple (but extreme) example, note that a rule that said only one pharmacist would henceforth be permitted to practice in the country would presumably pass the UK filter, although it would completely eliminate competition![71] Some of the earliest applied work on regulatory capture found the phenomenon in prospectively competitive markets such as interstate trucking.[72] Capture models point to other factors such as scope for administrative barriers to entry and the level of organisation of the rent-seeking group vs. the wider public when explaining successful rent-seeking behaviour.

---

[71] Indeed, only Questions 5 and 6 might be answered 'yes' in such a case.

[72] See for example Stigler (1971), pp.7-10, with a more detailed treatment in Rose (1985).

There is also a potentially serious practical problem with the UK filter concerning the role of market definition. We have noted earlier the mixing of questions about market characteristics with those about the nature of the proposed regulatory measure. Detailed questions about the market, and in particular market shares, presuppose some knowledge of the boundaries of the market. While there are well-established tools available for antitrust market definition, it is by no means clear that they are (or should be) applied at the early stage of regulatory impact assessment. A requirement to use these tools in all cases would be an onerous administrative burden, while not using them could imply that conclusions based on market characteristics are unsafe. One possible way to resolve this simultaneity problem would be for someone to produce a long list of candidate markets for all sectors likely to be affected by regulation. However, such an exercise might not be considered cost-effective in a large jurisdiction, never mind a small one.

The UK filter does not require a full-scale market definition exercise, but it avoids this by leaving considerable discretion with the policymaker applying the filter. This could increase the risk of regulatory capture in some instances (discussed in Section 3.1 above), and even where the filter is applied objectively, this element of the test may reduce certainty and increase the risk of error.

Finally, we have noted that the questions concerning market structure and rapidity of technological change (5 and 8) seem particularly sensitive to the counterfactual that is applied. In cases where regulation is already in place, it may already have affected market structure and innovation. For example, a regulatory measure that created a statutory monopoly would have a direct effect on market structure and could have the effect of slowing technological change in the relevant sector. If the filter were applied with a "no regulation" counterfactual, such effects should be picked up. However, in practice it might be applied with a "no change to regulation" counterfactual, and policymakers might

then answer Questions 5 and 8 based on the current (distorted) market conditions. This is akin to the "cellophane fallacy" described by Schmalensee (1987, pp.47-48).

3.3 Towards a better competition filter

It is clear that policymakers carrying out regulatory impact assessments need some way to determine when a detailed review of competition impact is required. Not reviewing competition impact would undermine the effectiveness of the RIA process and run the risk of serious consumer detriment. However, carrying out a detailed competition review of all proposed measures is unnecessary and impractical.

In this section we attempt to outline a practicable test to meet this need. Our emphasis is on capturing measures that run the highest risk of damage to competition; making the test easy to apply, even by staff without a great deal of technical economics training; and avoiding frequent reviews of non-problematical cases.

We start with the structure of the test, before turning to the sorts of questions that should be included in it. A sample test designed along these lines is shown in Figure 1 below. Note that this is only a sample test. The level of scrutiny to be given to measures at the filter stage as opposed to a full assessment and the balance between the specific competition assessment and the wider regulatory impact assessment process are likely to depend on the specific institutions and available administrative resources in a given jurisdiction.

Structure

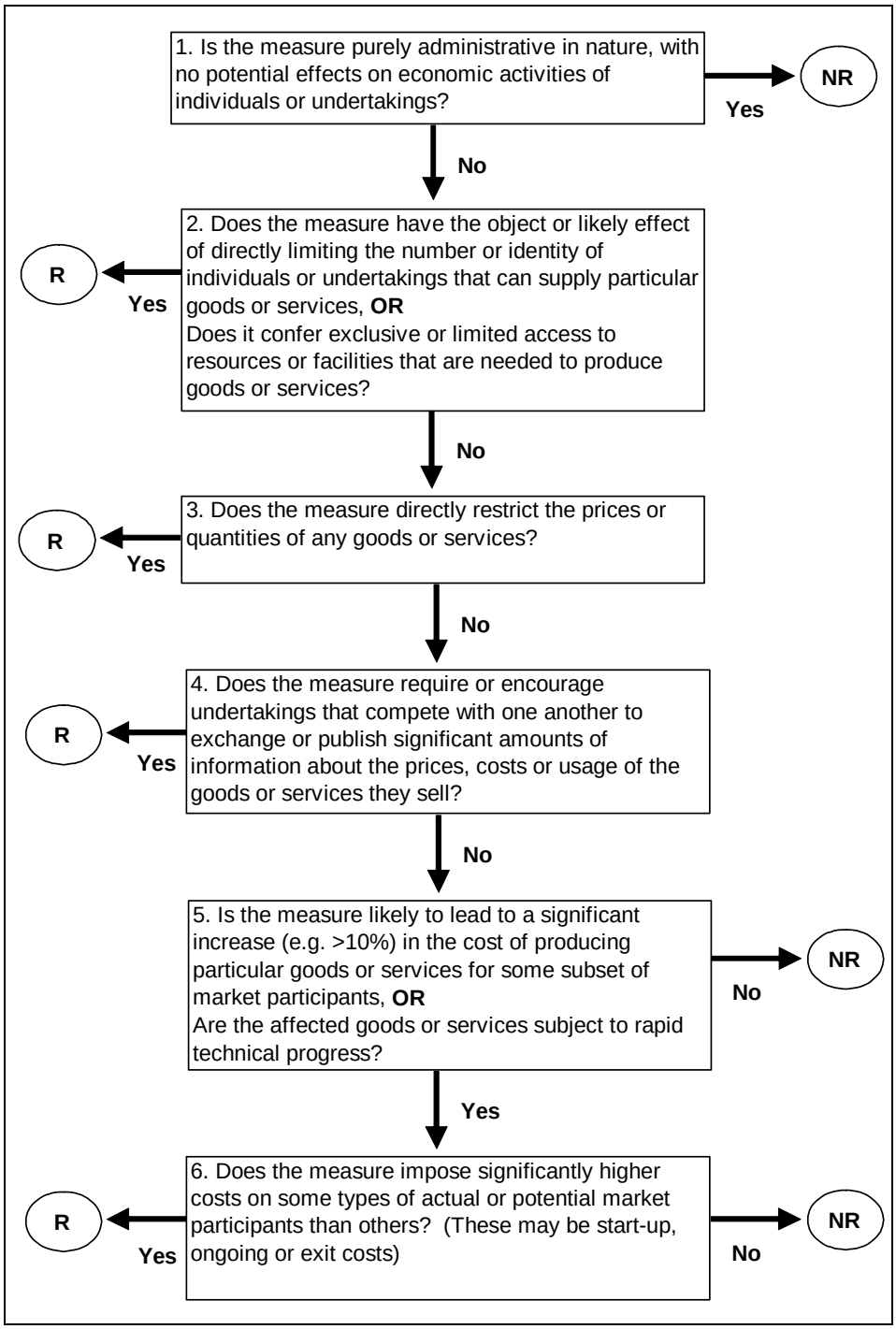
A "serial" filter structure is probably most appropriate, given that the test is likely to be carried out by personnel that are not specialists in the economics of competition. This approach reduces the exercise to a set of relatively simple steps, rather than a complex

multi-dimensional sum. If extra discrimination is desired and resources are available, additional steps can be added.

When examining the UK filter, we noted that a focus on *ex ante* concentrated markets seems to be counterproductive and makes the test very difficult for non-specialists to apply properly. A more practical alternative is to focus on the nature of the measure rather than features of the market. While the latter may be relevant to the impact of regulation under various (complex) circumstances, in practice these can only be taken into account in the context of a detailed investigation.[73] Removing most of the market-related analysis from the test should reduce the administrative burden associated with it significantly.

[73] If we were to include market structure variables in the test, this would imply that a market definition exercise would be required at the outset. Defining economic markets tends to require a significant commitment of specialist resources.

Figure 1: Example of a "serial" competition effects filter
(Key: "R"= review needed; "NR" = no review needed)



Questions to be included in the filter

Obviously, policymakers should not be required to spend scarce competition analysis resources on measures that have no discernable economic effects. Therefore, the first question in the filter could try to identify measures of this kind.

The next step is to ensure that the types of measures most likely to cause harm to competition trigger further scrutiny. From the discussion in Section 3.1 above, these include cases where the measure involves economic regulation, such as direct regulation of entry, prices and quality. Other measures that may have related effects are ones conferring exclusive or limited access to scarce resources.

With the increasing importance of informational regulation in many sectors, it might be appropriate to review measures that run the risk of facilitating tacit or explicit collusion, e.g. by requiring publication or exchange of strategically sensitive information.

Secondary features of concern should come into play only if the wider RIA exercise suggests they are present. These might include measures that have a high cost impact as a proportion of total sector costs, asymmetric effects on different types of firms in the sector, or those that involve regulation of innovative sectors. The precise formulation of these questions should depend upon the resources available for in-depth reviews.

4 False Negatives from the UK Competition Filter

In Chapter 3 we identified aspects of the structure and content of the UK competition filter ("UK filter")that lead to false negative results for important classes of potentially damaging regulation. These features are illustrated by case examples drawn from Ireland. However, it seems natural to ask whether similar shortcomings apply to the filter when applied in its native jurisdiction: the UK.

In this chapter, we extend the findings of Chapter 3 to a UK context by applying the UK filter to three selected UK measures that have previously attracted criticism from the Office of Fair Trading (OFT) on the basis that they appear to be damaging to competition. If measures such as these were to be proposed today (either *de novo* or with a view to retention), we suppose that they should trigger the filter, leading to a competition review. However, we find that the UK filter would probably clear these measures.

This is further evidence that the UK filter is prone to give false negative results when applied to important categories of measures that have a potential to harm competition.

4.1 Three examples of UK measures that the UK competition filter might exempt from detailed scrutiny

In this section, we consider how three hypothetical cases, based on existing or past UK measures, would fare under the UK filter. All of these measures appear to have the potential to harm competition if they were introduced today, and they are chosen to illustrate aspects of the UK filter's workings. Our approach is the same as that taken in Chapter 3: we are not asking whether these measures are justified or not, but rather exploring why the filter may clear measures that are potentially of concern.

Despite the apparent scope for harm to competition in each case, we shall see that the UK filter indicates that none of them requires a full competition review. The filter questions are repeated from Chapter 2 for ease of reference:

Table 7: UK competition filter questions

1. In the market(s) affected by the new regulation, does any firm have more than 10 per cent market share?
2. In the market(s) affected by the new regulation, does any firm have more than 20 per cent market share?
3. In the market(s) affected by the new regulation, do the largest three firms together have at least 50 per cent market share?
4. Would the costs of the regulation affect some firms substantially more than others?
5. Is the regulation likely to affect the market structure, changing the number or size of firms?
6. Would the regulation lead to higher set-up costs for new or potential firms compared with the costs for existing firms?
7. Would the regulation lead to higher ongoing costs for new or potential firms compared with the costs for existing firms?
8. Is the market characterised by rapid technological change?
9. Would the regulation restrict the ability of firms to choose the price, quality, range or location of their products?

Source: OFT (2002).

UK Example 1 – Taxis

In November 2003, the OFT published a market study titled "the regulation of licensed taxi and PHV services in the UK".[74] The study assembled evidence on the UK market and examined the effects of three types of regulation: limitations on the quantity of taxis, quality and safety rules and fare regulation. Our first example focuses on the first category: limits on the number of taxi licences issued.

[74] OFT (2003c).

After considering possible arguments for and against such limits, the OFT concluded that they are not justified.[75] Negative effects of the restrictions cited by the OFT included increases in waiting times and reductions in choice, safety and availability of taxis. Competition was found to be harmed by the limitation on market entry, both in the sense that applicants are delayed in entering the market and that entry is made more costly by the high premium on taxi licences. The study recommended that quantity restrictions be lifted by local authorities and suggested that the legislative provisions allowing them should be repealed.[76]

Suppose that new a new legislative measure permitting quantitative restrictions on taxis (or directly applying such restrictions) were proposed and subjected to a RIA. In light of the OFT's findings, it seems logical that the UK filter should require a competition review of such a measure. In Table 8 below, we apply the UK filter to a hypothetical quantity restriction on taxis.

[75] OFT (2003c), pp.23-44

[76] OFT (2003c), p.44.

Table 8: Application of UK filter to UK Example 1: *Restriction on the quantity of licensed taxis*		
Question(s)	**Likely Answer**	**Reasons**
1-3	No	Taxi markets, despite being local or regional in nature, appear to have very low concentration. For example, there were about 20,700 taxis and 2,000 private hire operators serving the London area in 2002.[77]
4	No	There are no obvious asymmetries in the cost of the regulation to different licensed operators.
5	Yes	The measure does affect market structure, via the number of taxis permitted.
6	Yes	New entrants would face either significant delays on a waiting list for a licence or a substantial cost to purchase one in the secondary market.
7	Yes	Ongoing costs do not seem to be affected in an asymmetric way by quantity restrictions.
8	No	Technological change does not appear to be rapid.
9	No	Although measures affecting price and quality of service are imposed by many local authorities, quantity limits *per se* do not include these types of restrictions.

Taxi markets, with most supply coming from a large number of small providers, are unlikely to be deemed concentrated, even if markets are defined on a local or regional basis (i.e. to be consistent with licensing boundaries). Thus the first three questions must be answered 'no'.

Once a firm is in the market, the playing field is level with respect to these measures, so question 4 must also be a 'no'. The answers to questions 5 and 6 are emphatically 'yes'; due to the strong asymmetry of treatment between incumbents and potential entrants, the

[77] OFT (2003c), Annex B, Table A.12.

market structure will be affected. However, even the 'yes' for question 5 might be avoided if the measure were being considered for retention (i.e. licensing was already restricted) rather than *de novo*.

However, since there are no asymmetric effects on ongoing costs (presumably post-entry), the technology is mature and other types of economic regulation are handled in separate restrictions, the answers to the other questions appear to be 'no'. In sum, only two questions require a 'yes' answer and quantity restrictions imposed on taxi markets do not trigger detailed competition scrutiny under the filter.

Like the first example given in Chapter 3, this measure passes the filter mainly because the relevant market is not found to be concentrated. Here too, the problem with the filter is that its structure is more prone to catch restrictions with a range of potentially harmful effects than measures that are harmful in a focused way.

UK Example 2 – Restrictions on professions complementary to dentistry

In 2003, following a study of the dentistry profession, the OFT concluded *inter alia* that a range of restrictions on "professions complementary to dentistry", such as dental hygienists, and on corporate dental bodies, had the effect of "limit[ing] choice, competition and the potential to develop and deliver innovative and better services."[78] The report recommended that complementary professions should be permitted more freedom to provide dentistry services to the public, and supported proposals to lift restrictions on corporate bodies serving this market.

[78] OFT (2003b), pp.6-7. The relevant statute was the Dentists Act, 1984.

If restrictions such as these were to be proposed by a policymaker, how would they fare under the filter? Given the OFT's stated concerns about the existing restrictions, presumably the filter should indicate that such measures have the potential to harm competition. To find out whether it would, we again use published information to answer the nine questions (see Table 9 below).

Table 9: Application of UK filter to UK Example 2:
Restrictions on professions complementary to dentistry and on corporate form of dentistry providers

Question(s)	Likely Answer	Reasons
1-3	No	According to the OFT, there were about 11,000 dental practices in the UK in 2001.[79] It seems unlikely that the relevant market(s) , whatever their geographical extent, would exceed the concentration levels set out in these questions.
4	No	There are no obvious asymmetries in the cost of the regulation to different providers.
5	Yes	To the extent that limits on complementary professions restricted entry, market structure might be affected.
6	Yes	Potential market entrants from complementary disciplines would face a different start-up cost from entrants from within the dentistry profession.[80]
7	Yes	To the extent that some providers were permitted to employ a corporate structure and others were not (e.g. under the provisions of the Dentists Act, 1984),[81] the two types of firms could experience different ongoing costs. In addition, restrictions on the work that complementary professions may carry out could affect the relative costs of practices depending on staff mix.
8	No	Technological change does not appear to be rapid.
9	No	Although measures affecting price and quality of dentistry services may be imposed by other regulatory measures, restrictions on complementary professions or on incorporation *per se* do not include these types of regulation.

[79] OFT (2003b), p.24.

[80] As in the case of Example 1, this 'yes' depends upon use of a no-regulation counterfactual.

[81] OFT (2003b), para 1.28.

With 11,000 practices serving the market, it seems unlikely that UK dentistry services could exceed the UK filter's concentration thresholds, so the answer to Questions 1-3 would appear to be 'no'. Measures restricting right to practice and corporate bodies would arguably give rise to asymmetries between potential entrants and existing firms (i.e. where costs would depend upon a given practice's corporate structure or the qualifications of its staff). However, these issues, affecting Questions 5 through 7, could yield at most three 'yes' answers in the filter. Since Questions 4, 8 and 9 appear to yield 'no' answers, the filter would not be triggered.

Note that we obtain this result despite bundling together a set of measures that could, in principle, be examined separately. In general, applying the UK filter to a set of measures will give a higher likelihood of five 'yes' answers than applying it to a measure in isolation.

This example suggests that even measures (or sets of measures) with relatively broad effects may not be captured by the filter if the markets to which they are applied do not trigger the concentration questions.

UK Example 3: Pharmacy entry restrictions

Since 1987, regulations have restricted the opening of community pharmacies in the UK.[82] We will focus on the version of these restrictions contained in the National Health Service (Pharmaceutical Services) Regulations, 1992. Among other restrictions, applications to open new premises were accepted only when it was deemed

[82] OFT (2003a), p.20.

> *…necessary or desirable to grant the application in order to secure, in the neighbourhood in which the premises from which the applicant intends to provide the services are located, the adequate provision, by persons included in the list, of the services, or some of the services, specified in the application.*[83]

An OFT report on community pharmacy services published in 2003 found that entry restrictions on community pharmacies "acted to impede entry and expansion by pharmacies that offer consumers lower prices, more convenient opening times, or valued and innovative services. Moreover, by limiting the numbers and location of pharmacies in a local area, the regulations have restricted competition between pharmacies, in terms of both prices and quality of service."[84] The report recommended that these restrictions be ended.[85] After a review by the government, legislation was enacted providing exemptions from entry restrictions subject to a range of conditions.[86] As before, we suppose that if measures like these were to be proposed in the future, their potential competition effects ought to receive detailed scrutiny.

As before, we apply the UK filter to the control of entry regulations applied to community pharmacies. In contrast to our first example, the UK market may be concentrated enough to trigger at least two of the relevant questions (see table below). According to the OFT, the largest pharmacy chain, Lloydspharmacy, had a 10.9% share of total UK contractor pharmacy outlets in 2002, and the corresponding share of the top two firms was 21.4% of

---

[83] Regulation 4(4) of the National Health Service (Pharmaceutical Services) Regulations, 1992.

[84] OFT (2003a), p.3. The regulations cited were the National Health Service (Pharmaceutical Services) Regulations, 1992 covering England and Wales; the National Health Service (Pharmaceutical Services) (Scotland) Regulations, 1995; and the Pharmaceutical Services Regulations (Northern Ireland), 1997.

[85] OFT (2003a), p.6.

[86] The relevant measure for England and Wales was the NHS (Pharmaceutical Services) Regulations 2005

outlets.[87] This suggests that Questions 1 and 2 would merit a 'yes'. However, the share of the top three firms was only 27.8% (i.e. below 50%), suggesting that Questions 3 would be answered 'no'.

Table 10: Application of UK filter to UK Example 3:
Control of entry regulations on pharmacies

Question(s)	Likely Answer	Reasons
1-2	Yes	The largest firm had over 10% of outlets in 2002 (totalling 1,321) and the largest two firms had 21.4%.
3	Yes/No	The largest three firms had only 27.8% of outlets. However, if relevant geographical markets was taken to be local, the answer could be 'yes' in some markets, but 'no' in others.
4	No	Costs arising from restrictions on the quantity of pharmacies do not appear to have an asymmetric incidence on existing firms.
5	Yes	The measure does affect market structure, via the number of pharmacies permitted to enter.[88]
6	Yes	Start-up costs do appear to be affected significantly by the entry restrictions.
7	No	Ongoing costs do not seem to be affected in an asymmetric way by entry restrictions.
8	No	Technological change would probably not be considered rapid, as per Example 2 in Chapter 3.
9	No	Although measures affecting price and quality of pharmacy services may be imposed by other regulatory measures (e.g. via the NHS dispensing contract), entry restrictions *per se* do not include these types of regulation.

[87] OFT (2003a), Table 2.1.

[88] As in the case of Example 1, this 'yes' depends upon use of a no-regulation counterfactual.

Of course, in focusing on national market share figures, we are implicitly assuming that the relevant market is national in scope. That is the easiest option to apply, in terms of readily-available data, and as we shall see, assuming a national market would probably result in the measure passing the UK filter.

Of the other questions, perhaps only the fifth and sixth (concerning market effects on market structure and start-up costs) would merit a 'yes'. In particular, it seems unlikely that pharmacies would be considered technologically innovative, and the entry restrictions, considered in isolation, contain no price or quality regulatory measures.

Assuming a national market, we would thus obtain only four 'yes' answers, and if pharmacy entry restrictions were to be included in proposed legislation today, they would not trigger the UK filter.

However, note that the UK filter is close to obtaining five 'yes' answers, and there is an obvious way it could do so: if markets were instead defined on a local basis. Some local markets would doubtless be more concentrated than the average and thus lead to a 'yes' for Question 3. Of course, other markets would be less concentrated than the average, leading to a 'no'. A strict application of the filter might suggest that a full competition analysis would be required for the measure's effects in some areas, but not others. Does this provide any comfort as to the effectiveness of the filter? Perhaps, but it also emphasises the vulnerability of the filter to ambiguities and its reliance on the market analysis skills of policymakers.

Recall that a finding of national markets would allow the use of data that are easy to obtain, avoiding the need to collect data on local markets. It would also lead to a "negative" result from the filter that would minimise the remaining work to complete the regulatory impact assessment process. Moreover, some features of the market could be

seen to support a finding that it is national in scope; for example, pricing policies for many of the chains are set nationally[89] and payments for NHS dispensing contracts (an important component of sales) are based on a common national scale.

The inclusion of market characteristics in the UK filter requires that geographic and product markets be defined. This element of discretion, combined with scarcity of specialist analytical resources and an incentive to shade the filter towards a negative result to minimise the administrative burden of RIA, risks encouraging "discretionary" false negatives in addition to the ones generated by the structure and content of the filter.

[89] OFT (2003a), p.32.

Summary of examples

The table below summarises the filter's results for the three examples given above.

Table 11: Summary of UK examples (✓ = yes, - = no)			
Question	**Example 1** *Restriction on the quantity of licensed taxis*	**Example 2** *Restrictions on professions complementary to dentistry and on corporate form of dentistry providers*	**Example 3** *Control of entry regulations on pharmacies*
1	-	-	✓
2	-	-	✓
3	-	-	(- / ✓)
4	-	-	-
5	✓	✓	✓
6	✓	✓	✓
7	-	✓	-
8	-	-	-
9	-	-	-
Overall result	**Pass** (Yes = 2)	**Pass** (Yes = 3)	**Ambiguous** (Yes = 4-5)

Source: analysis by the authors.

In Example 1, we saw that a focused measure affecting a non-concentrated market is very unlikely to trigger the UK filter. Example 2 extends that lesson to measures with somewhat broader effects, but which are still applied to markets that would be deemed un-concentrated, at least *ex ante*. Our third example highlights some potential problems arising from the inclusion of market characteristics in the filter; in particular, the heightened breadth of policymaker discretion, combined with an additional analytical burden. For completeness, we note that the examples in Chapter 3 also include another

class of measure that may be prone to false negative results: one in a concentrated market, but with focused effects (e.g. a price or entry control examined in isolation).

4.2 Conclusions

We have discussed three examples of UK regulatory measures that have been criticised by the OFT, yet might not trigger the UK filter if they were introduced today. These examples discussed illustrate the same shortcomings of the filter identified in Chapter 3, and we summarise them below.

Structural problems arise from the use of a "parallel" test in which disparate questions are applied simultaneously and the results are aggregated, with all answers being accorded equal weight. This approach makes the test prone to failure when confronted with measures that derive their harmful impact on competition from a focused effect, rather than a broad set of effects. Measures prohibiting entry in prospectively competitive markets are the obvious examples here.

This structure also creates an incentive for advocates of a proposal to press for it to be introduced, or at least examined, on a piecemeal basis. Unless the effects of proposed measures are assessed in conjunction with other existing or planned regulations, this incentive may encourage further false negative results.

There are also problems with the **content** of the filter. The nine questions mix analysis of regulatory *measures* with analysis of regulated *markets*. Among the latter, the questions on the extent of concentration seem particularly problematic.

First, high concentration is not a reliable indicator for the vulnerability of a market to distortion by regulatory measures. Imposing entry or price regulation on an un-concentrated, highly competitive market might pass the filter, but it could be at least as

damaging to welfare as imposing the same sort of regulation on a market that is concentrated *ex ante*.

Second, the need to answer such questions implies that the policymaker must form a view as to the boundaries of economic markets to which the measure will apply. This is a complex task requiring specialised economic analysis. Given that concentration is a poor indicator of market vulnerability, requiring this type of analysis seems administratively burdensome.

PART II: EMPIRICAL ANALYSES OF REGULATORY IMPACT

In Chapter 2, we suggested that one way to improve the administrative practicability of RIA in a small jurisdiction, while maintaining analytical rigour, is to draw upon the experiences of other countries that have previously employed the measure under consideration. In Part II of this thesis, we discuss three empirical applications of this approach.

In Chapter 5, we show that international data can be used to quantify some likely effects of imposing mobile number portability (MNP), a regulatory measure intended to reduce customer switching costs. Past *ex ante* cost-benefit analyses of MNP regulations tended to include an element of guesswork, particularly concerning the likely effects of the measure on retail prices and customers' propensity to switch. Our analysis shows that it is now possible to estimate such parameters using international data. In line with theory, we find that prices tend to fall and switching tends to rise in markets where a sufficiently high "quality" of MNP is employed.

Chapter 6 raises some cautionary points. While finding shortcuts is often necessary – and even desirable – to operate an efficient RIA process, there are also risks of over-simplification. Staying with the example of mobile telephony markets, we examine whether the commonly-used metric of Average Revenue Per User (ARPU) really means what some regulators say it does about competitive conditions. We find that inclusion of other factors (especially proxies for demand) in a model of ARPU removes the apparent association between ARPU and concentration, suggesting that the use of this metric as an indicator of intensity of competition may be unsound.

Chapter 7 looks at how far international survey evidence of regulatory (and other) constraints on small firms conveys useful information about likely post-entry performance of new micro businesses. We find that existing cross-country data does have some explanatory power, but the evidence on the relative importance of different constraints is less conclusive. We suggest changes to the design of small business surveys to better isolate the effects of the various possible constraints on performance.

Note: quantitative information presented in Part 2 is generally rounded to three significant digits.

5 Measuring the Benefits of Mobile Number Portability

Increasing numbers of countries require mobile telephone network operators to offer mobile number portability (MNP). This facility allows customers who wish to switch mobile operator to keep the mobile numbers originally assigned to them, avoiding the costs of switching to new numbers.

Since MNP regulation was first mooted, policymakers have asked whether it can produce positive net benefits. *Ex ante* evaluations of MNP carried out in several countries have produced detailed estimates of expected costs and direct benefits (e.g. the savings accruing to customers from lower switching costs). While researchers have suggested MNP should have a range of potentially important effects, such as strengthened competition and reduced prices (see Buehler, Dewenter and Haucap (2006) for a recent discussion), few attempts have been made to quantify them *ex post*.

The staggered introduction of MNP internationally provides a useful natural experiment. In this chapter, we use econometric analysis of international time-series cross-section data to estimate the average treatment effects of MNP on retail prices and switching by customers. The dataset constructed for this purpose includes information from up to 38 countries for 22 quarters (1Q 1999 through 2Q 2004).

We find that the quality of MNP, as proxied by the target maximum porting time, helps explain its impact on switching and average prices. For countries in our sample that required porting to be completed in five or fewer days, MNP was associated with increased customer switching and lower prices. The sub-sample of countries with less stringent porting time standards experienced no significant churn or revenue effects.

The costs associated with the MNP service depend upon the technology used to deliver it (Buehler, Dewenter and Haucap, 2006). The technology, in turn, determines the "quality" of MNP, including dimensions such as porting time and reliability. Previous research, e.g. Gans, King and Woodbridge (2001), has emphasised the importance that the choice of number portability technology has in determining the likely effects of the measure. Our results provide empirical support for this view. Jurisdictions conducting *ex ante* assessments of MNP in the future should consider the likely trade-off between achieving positive market outcomes and cost of implementation.

Section 5.1 of this chapter provides a brief classification of the potential benefits of MNP and refers to some previous research, including both *ex ante* cost-benefit studies and other empirical research. In Section 5.2, we ask what effects MNP should be expected to have on consumer switching behaviour and prices. The dataset constructed for this study is described in Section 5.3, along with some descriptive statistics. Sections 5.4 and 5.5 set out econometric models of switching and retail prices, respectively, and Section 5.6 discusses our conclusions and suggestions for future research.

5.1 Potential benefits and costs of mobile number portability

To provide context for the empirical analysis that follows, in this section we briefly review some relevant empirical research. This consists of *ex ante* cost benefit analyses conducted on MNP by regulators and a modest number of *ex post* empirical studies. Existing theoretical research on mobile number portability was recently surveyed in Buehler and Haucap (2004), but to clarify terminology used in the remainder of the section, it is worth restating the standard classification of number portability benefits.

Classification of benefits

A commonly-used approach to analysing the likely costs and benefits of MNP divides the measure's potential benefits into three types, the first of which is divided into two subtypes. They are summarised briefly below:[90]

> **Type 1** benefits are those obtained directly by customers who switch. These are divided by beneficiary group:
>
> > **Type 1A** benefits to customers that will switch whether or not MNP is available;
> >
> > **Type 1B** benefits to customers that will switch only if they can avail of MNP;
>
> **Type 2** benefits are those that are obtained by all mobile telephony customers (e.g. efficiency gains and price reductions due to strengthening of competition); and
>
> **Type 3** benefits are obtained by those making calls to ported numbers.

Past *ex ante* evaluations have proceeded on the basis that MNP should be expected to provide net welfare gains if the sum of these benefits exceeds the cost of network investments, process changes and operating expenses incurred to make mobile numbers portable.[91] However, they have tended to focus on the more empirically tractable Type 1 and Type 3 benefits, giving less emphasis to Type 2 benefits. In the next section we review some of the results of these *ex ante* evaluations.

---

[90] This framework was originally devised by NERA for the UK regulator OFTEL in a study of geographical number portability: Monopolies and Mergers Commission (1995), pp.58-59. See Oftel (1997) for an early application to mobile number portability.

[91] Some estimates of MNP's costs based on U.S. data are provided in Lenard and Mast (2003), pp.11-20.

Ex-ante Cost-benefit Analyses

Full mobile number portability (MNP) was first employed in Singapore in 1997, and since then many countries have introduced this form of regulation. Several cost-benefit analyses (CBAs) are available in published form, notably Oftel (1997) for the UK, NERA/Smith (1998) for Hong Kong, and Ovum (2000) for Ireland. In Table 12 below, we summarise the estimated benefits per customer by type from each of these studies.

Table 12: Predictions from three *ex ante* assessments of MNP

Country	**UK**	**Hong Kong**	**Ireland**
Base year	1997	1998	2000
Expected benefits per subscriber *Present value (in USD) of ten year impact divided by subscribers in base year*			
Type 1	28 - 81	39 - 71	78
Type 2	n/a	1	26
Type 3	1 - 5	1 - 3	5

Sources: analysis of estimates in Oftel (1997), NERA/Smith (1998) and Ovum (2000). Exchange rates are base year figures from IMF International Financial Statistics.

Type 2 benefits were viewed as difficult to estimate, and since Type 1 benefits were by themselves expected to be sufficiently high to justify the intervention, Type 2 benefits were either not quantified or subject to only simple scenario analysis. For example, in the CBA for the Irish market, Ovum assumed that MNP would lead to a 3% fall in retail post-pay mobile telephony prices.[92] Sensitivity analysis was carried out for reductions of 1% and 5%. Ovum also noted that there might be benefits from cost efficiencies or greater innovation, but these were not modelled.

The high level of Type 1 benefits incorporated in *ex ante* assessments was not unreasonable given the high level of switching in mobile markets prior to MNP. High

[92] Ovum (2000), pp.12-13.

estimates of Type 1A benefits require only that switchers use MNP rather than giving up their numbers when moving. However, more recent evidence suggests that use of MNP by switchers is not as common as some *ex ante* analyses expected. We discuss some of this evidence in the next sub-section.

Large *ex ante* estimates of Type 1B benefits relied on predictions that the rate of switching, proxied by churn, would increase significantly. For example, in the Hong Kong CBA, scenarios were examined allowing for increases of 5-15% in the churn rate following introduction of MNP.[93] Oftel (1997) includes an assumption in its model that, in the first ten years after implementation of MNP in the UK, over 50% more customers would switch operators than would have switched in the absence of the intervention.[94]

Other empirical research on the effects of MNP

The main *ex post* empirical work on MNP to date has focused on the propensity of those switching mobile provider to use MNP. This is particularly relevant to the size of Type 1A benefits as discussed above.

As part of a wider study of switching costs for the UK Office of Fair Trading, NERA (2003) examined the usage of MNP for inter-operator switching in UK mobile telephony markets. They found that in the first two years after MNP was introduced, the usage of MNP was very limited for residential customers, with only 12% of customers that switched operator taking up the portability option. This is far lower than the rate predicted in *ex ante* assessments. However, half of businesses who changed numbers in this period

[93] NERA/Smith (1998), p.66.

[94] To be precise, the analysis assumed that there would be half the number of Type B customers (who switched due to MNP) as Type A customers (who would have switched anyway). Oftel (1997), Section 3.3 and appendices.

ported at least some of their numbers. NERA suggested that the difficulty of using MNP during the first years after implementation may explain its unpopularity: porting a number originally took an average of 25 days. When the delivery time was reduced to five days on average, take-up increased to about 18% for residential customers and 80% for businesses.[95]

Looking beyond the propensity of switchers to use MNP, there has been little previous empirical work on the broader effects of MNP regulation. Ovum (2005) examined the experience of MNP in six countries that have implemented it: Australia, Germany, Hong Kong, Ireland, the Netherlands and the UK. Several of their findings are relevant to this study:

- Usage of MNP can fall significantly if the time it takes to change operator ("porting time") is too long. The authors suggest that two days is a practical upper limit. However, very short porting times do not necessarily increase demand for MNP.
- High end-user charges for MNP can also deter usage of the facility. Lower charges, which the authors suggest are levels of less than 20% of monthly average revenue per user, do not seem to be a "major deterrent to usage".[96] However, zero charges do not seem to increase demand beyond the levels associated with low charges.
- In jurisdictions with MNP, the extent to which switching customers use it varies widely and tends to increase over time. Ovum concludes that "adoption rates for

[95] NERA (2003), pp.37-39.

[96] Ovum (2005), p.1.

> Mobile Number Portability have been disappointing in all countries other than Hong Kong," but they point to an improving trend in some of the more recent implementations such as Ireland and Australia.[97]

There has also been a limited amount of academic research on individual markets. Below we cite two concerning MNP and one on number portability in a related market.

Lee, Kim and Park (2004) used contingent valuation techniques to estimate the prospective demand for MNP in South Korea. They found that the average South Korean mobile user was willing to pay an average of 3.24% of his or her monthly bill for a mobile number portability option. Willingness to pay (WTP) showed a strong positive association with income, awareness of MNP, and intention to switch. The authors also found that WTP varied significantly depending upon a user's network operator: the figure was lower for customers of the incumbent operator than those using either of the alternative operators. Other demographic variables such as age, gender and occupation were not found to be significant.

A recent *ex post* study of MNP's effects also focuses on South Korea. Kim (2005) estimated switching costs for customers of two of the country's mobile network operators by applying a random utility model to cross-sectional subscriber-level microdata. The paper compared switching costs calculated using samples before and after MNP, and differences between these estimates were attributed to MNP. Controls included firm-specific dummy variables, prices, non-price network attributes and customer characteristics. The paper estimated that MNP reduced average switching costs in South

[97] Ovum (2005), p.5.

Korea by more than 35%.[98] Data reported in the paper indicates that there was significantly more switching after MNP was introduced, at least among customers of the largest operators.[99] Service fees maintained a downward trend of about 7% per annum from 2002-2005, with no obvious change in relative or absolute prices at the point MNP was introduced for the two largest operators (July 2003).[100] Per-minute prices remained broadly unchanged over the period.[101]

Viard (forthcoming) examined the effect of number portability on prices in the US market for toll-free calls. This service is different from mobile telephony, but it is similar in some respects (for example, it is subject to high rates of growth).[102] Estimating price regressions on data from 219 AT&T virtual private network contracts, he found that introduction of number portability was associated with price reductions of 4.4%. A control group of contracts containing no toll-free services showed no relationship between prices and the introduction of number portability. Viard interpreted the results as evidence of an inverse relationship between switching costs and competition in this market: "despite rapid growth in the market, the firms' incentive to exploit their existing 'locked in' users was greater than their incentive to 'lock in' new customers."[103]

---

[98] Kim (2005), p.16.

[99] *Ibid*, Table 2.

[100] *Ibid*, p.11.

[101] *Ibid*, Figure 5.

[102] Note, however, that there are also important differences between mobile telephony and toll-free calls markets; in particular, mobile operators may be able to price discriminate between new and existing users. NERA (2003) noted that handset subsidies in effect involve lower prices for new customers than for existing ones; pp.30-31.

[103] Viard (2004), p.25.

Empirical research on wider effects of regulation on mobile telephony markets

There is some published empirical research into the effect of regulatory measures on mobile telephony prices, much of it prepared in the context of US regulatory proceedings. For example, using panel data for US mobile telephony markets, Hausman (1995) and Hausman and Kuersteiner (2004) found that regulating mobile tariffs had the perverse effect of significantly increasing retail prices (in comparison to not regulating) and that regulation was the strongest determinant of prices among the variables tested. These models explain prices for baskets of calls purchased by "light", "medium" and "heavy" users of cellular services by regressing them on population, commuting time, income and type of spectrum used by the relevant operator.

Shew (1994), in another study of US cellular operators, used panel data regressions to test the effects of different types of price control measures and found little evidence that regulation affected prices. His results indicated that licensing more that one operator reduced prices in the relevant markets, and he identified other variables that affected mobile telephony prices, including income, commuting time, population density, cost of living, and the length of time services had been offered in a market.

Hazlett and Muñoz (2004) estimated a system of demand and mark-up pricing equations to examine the effects of mobile spectrum allocation policies. Of relevance to our analysis, they found mobile market concentration (using the HHI[104]) to be positively associated with prices (proxied by average revenue per minute).[105]

---

[104] Herfindahl-Hirschman Index: the sum of squared market shares.

[105] Hazlett and Muñoz (2004), p.15.

Wallsten (2001) explored the impact of privatisation and regulation on indicators of telecommunications market performance in 30 developing countries, using annual data from 1984-1997. The study used fixed effects regressions to estimate the effects of a set of regulatory, competition and macroeconomic variables on a set of performance measures. Of relevance to mobile telephony, Wallsten found a positive association between the number of mobile network operators in a country (beyond those owned by the incumbent) and fixed line penetration, connection capacity and the number of payphones, while the number of such operators had a negative relationship to fixed line local call prices.

5.2 Likely effects of MNP on switching and prices

In this section, we outline the main effects that economic theory suggests MNP should have on switching propensity and retail prices.

MNP and consumer switching

Significant numbers of customers switch operators at some point after their initial acquisition of a mobile subscription. There are likely to be many reasons for such switching, e.g. changes in individual demand patterns, service innovation, learning by customers about the fit between their pattern of demand and operator offerings, and changing price and quality propositions.

To the extent that the component of switching cost associated with changing one's telephone number is high enough to deter some customers from switching operator when they might otherwise have done so, MNP should yield a positive change in the conditional probability of switching (holding other variables constant). This effect might be offset in whole or in part by operators' reactions, e.g. if operators respond to MNP by reducing

price dispersion. Nevertheless, MNP should have at least a weakly positive effect on switching.

MNP and retail prices

The net effect of MNP on retail prices is in principle indeterminate. Empirically, it is likely to depend upon the interplay of three groups of effects:

- Pass-through of costs associated with the facility (increase in prices);
- Effects on competition (probably a decrease in prices); and
- Loss of customer information (increase in prices).

First, and most obviously, the implementation of MNP imposes costs on all operators employing it. Depending upon the extent of competition in a given national market, these costs are likely to be (at least partly) passed on to consumers and thereby lead to increased prices. Some argue that this is likely to be the main effect of number portability, and hence that mandating it through regulation will lead to a net reduction in welfare; see, for example, Ellig (2005, p.29). Aoki and Small (1999) also address the welfare impact of switching cost reductions due to number portability. They identify cases in which switching costs reductions provided by number portability (e.g. reducing the need to purchase complementary goods such as stationery) could be offset by higher marginal costs of providing call services, leaving consumers with lower surplus.

Beyond the simple effect of increased direct costs from implementation of MNP, theory is less definite about the effect of decreased switching costs on prices.

There are reasons to think that MNP might lead to lower prices. A survey by Klemperer (1995) on the effects of consumer switching costs on competition and concludes that "switching costs generally raise prices and create deadweight losses of the usual kind in a

closed oligopoly."[106] He further argues that such costs may discourage entry and reduce incentives to provide varied products, and he takes the view that policymakers should encourage activities that reduce consumer switching costs.

A related argument is that reducing switching costs also have the effect of lowering search costs. Switching may be an important means of search in mobile telephony markets, given the complexity of mobile telephony demand, services and tariffs. Lower switching costs might provide an incentive for customers to search more intensely, reducing the informational rents possessed by mobile operators. By leading to intensified search, MNP might reduce operators' market power and induce them to reduce prices.

However, there are exceptions to the favourable theoretical view of switching cost reductions. For example, Padilla (1995) suggests that high switching costs could make it more difficult to sustain tacit collusion, because by weakening competition in future periods such costs make it more difficult to enforce tacit agreements. This implies that a switching cost reduction might have the effect of facilitating tacit collusion (and thus potentially lead to a price rise).

Indeterminacy also arises on the search cost side. Samuelson and Zhang (1992) set out a model in which a fall in search costs can result in higher prices. They point out that a reduction in search costs may increase total demand, and hence prices, potentially offsetting the impact of the change on the marginal demand for each firm's services. It seems that we cannot be certain that MNP will reduce prices, even if the balance of arguments point that way.

---

[106] Klemperer (1995), p.536.

Finally, there is an informational channel through which MNP may lead to increases in at least one component of mobile telephony prices. Depending upon how MNP is implemented, it may reduce the tariff information available to both fixed and mobile customers wishing to make calls to mobile numbers. This effect is discussed in Buehler and Haucap (2004) and Gans and King (2000). Without MNP, mobile number prefixes provide a simple means for callers to identify which mobile network is going to terminate a given call. Fixed and mobile operators can therefore use these prefixes to differentiate retail call tariffs by terminating network. Introducing MNP necessarily removes the simple mapping between number prefixes and networks, so it may reduce the scope for providing retail tariffs specific to each terminating operator. Such a decrease in transparency could lead to higher prices for call termination.

Buehler and Haucap (2004) suggest that this informational problem might be avoided if an alternative mechanism were put in place to allow callers to identify the terminating mobile network for each call; for example, an audible tone or message could be included at the start of each call. However, even if such a mechanism were used, its direct cost and any loss in consumer convenience would need to be set against the improvements in tariff transparency achieved by it.

A second means of mitigating MNP's effect on tariff transparency probably has more practical relevance (at least in the short term). Many regulators have moved to regulate termination rates of mobile operators, taking the view that operators in a "calling party pays" regime possess market power over termination. In the presence of such regulation, mobile network operators are unlikely to be able to exploit the loss of transparency associated with MNP.

5.3 Data employed

We have constructed an unbalanced time-series cross-section dataset that includes most of the OECD and a selection of developing countries. It is based principally on the Merrill Lynch Global Wireless Matrix (Merrill Lynch, 2004).[107]

Although this source provides some data on 46 countries, there are many gaps. Also, we found that data for three countries, China, the Czech Republic and South Korea, contained implausibly large fluctuations in reported subscriber numbers. As a result, these countries were excluded from the dataset. The available panel includes data on 38 countries (for churn modelling) and 37 countries (for price modelling). See Table 21 in Annex 5.1 for details of the countries and the sample coverage.

The data are quarterly, running for up to 22 quarters from 1Q 1999 through 2Q 2004, and we omit the first two quarters to allow use of differenced and lagged variables. Table 13 below lists the variables and provides summary statistics; individual observations are for country *i* and quarter *t* in each case. Further information on some of the variables is provided in Annex 5.1.

[107] We are grateful to Elaine Pryor at Merrill Lynch, for permission to use these data.

Table 13: MNP dataset variable descriptions, sources and summary statistics

			Churn model		Price model	
Variable	**Description**	**Source**	**Mean**	**St Dev**	**Mean**	**St Dev**
MNP_{it}	= 1 if mobile number portability in place at any time in quarter *t*	See Table 21 in Annex 5.1	0.285	0.452	0.240	0.428
$MNPtime_{it}$	Target maximum single line porting period (days)	*Ibid.*	1.81	4.18	1.69	4.22
$MNP5d_{it}$	If MNP = 1 and *MNPtime* <=5 then 1, else 0	*Ibid.*	0.175	0.381	0.128	0.334
$MNP6p_{it}$	If MNP = 1 and *MNPtime* >5 then 1, else 0	*Ibid.*	0.109	0.312	0.112	0.316
$RMNP_{it}$	If MNP = 1, then (1/*MNPTime*), else 0	*Ibid.*	0.390	1.94	0.382	1.96
$CHURN_{it}$	Monthly number of disconnections from a network expressed as % of MNO's avg. subscriber base in the same month. Quarterly avg. of monthly rates.	Weighted avg. of individual MNOs' data from ML	0.0205	0.0102		
DEN_{it}	Cellular density: mobile users as a share of population	Analysis of ML	0.534	0.298		
OPS_{it}	Number of MNOs in country *i*	Analysis of ML	3.76	1.23	3.72	1.21
$RGDPPC_{it}$	Real GDP per capita (US$)	See Annex 5.1	17,400	12,300	17,100	12,100
RPM_{it}	Average real revenue per minute for MNOs in country *i* (US$)[108]	Weighted avg. of individual MNOs' data from ML			0.198	0.0794
$TOTMIN_{it}$	Monthly average minutes of mobile telephony traffic in country *i* (millions)	Analysis of ML			3,710	10,000
$PDNST_{it}$	Population density: population per Km^2	World Bank WDI (2004)			126	144
HHI_{it}	Herfindahl-Hirschman Index: Sum of the squares of the market shares (users) of all MNOs in country *i*	Analysis of ML			3,790	976
$CR1_{it}$	The top MNO's share of total users	Analysis of ML			0.477	0.116

Notes: MNO is an abbreviation for "mobile network operator". Merrill Lynch (2004) is referred to as ML.

[108] This is rebased to year 2000 prices using GDP deflators and it excludes revenue from data services.

5.4 Modelling the effect of MNP on switching

In this section, we define and estimate two econometric models of switching frequency, including proxy variables to capture the effect of MNP.

The switching variable

The ideal measure of switching for our purposes would directly identify flows of customers between operators, but such data are generally not put in the public domain. The best available proxy is **churn**, a metric based on the number of disconnections from each network as a proportion of the average number of network users in a given period. While inter-operator switching does feed into churn, the churn rate is not a pure measure of switching. Subscribers that leave a network without joining another one, for whatever reason, also appear as churn, as do customers on prepaid tariff packages that do not use their phones for a specified period.

Because churn is a proportion (but does not take values of zero or one in our sample), we apply a logistic transformation to the data before using it as a dependent variable:

$$LGTCHURN_{it} = \ln\left(\frac{CHURN_{it}}{1 - CHURN_{it}}\right) \qquad (1)$$

Explanatory variables

Switching propensity should be positively related to the presence or absence of MNP and to the quality of the MNP service, insofar as the service reduces consumer switching costs. However, we have no theoretical prior as to the functional form of the relationship. To

allow for a range of possibilities, we test two alternative proxies for MNP, both based on the target maximum porting time (*MNPTM*) in force in a given country.[109]

The first is a threshold approach, distinguishing between countries with a *MNPTM* of 5 days or less (for which *MNP5D* is set to 1) and those with 6 days or more (for which *MNP6P* is set to 1). Both variables are set to zero for all other cases. This divides the observations where MNP was in place into two roughly equal parts along the quality dimension. The second MNP proxy, *MNPR* is equal to the reciprocal of *MNPTM* for observations with MNP and to zero for those without the service.

In the remainder of this section, we include some descriptive statistics to illustrate the key bivariate relationships in our data.

A comparison of averages suggests that countries with "high quality" MNP had slightly higher churn than those without MNP, but those with "low quality" MNP had slightly lower churn (see Table 14). However, both MNP variables are positively correlated churn following a logistic transformation, as shown in the correlation matrix for our churn sample (Table 22 in Annex 5.1 below).

Table 14: Relationship between churn rates and mobile number portability

Case	Sample mean quarterly churn
No MNP	0.0203
MNP delivery time <= 5 days	0.0218
MNP delivery time 6+ days	0.0198

Source: see Table 13 above.

[109] Data on actual, rather than target, porting times would probably be a better measure of quality. Unfortunately, these data are not made public in most countries.

In a regression analysis, we expect coefficients on both *MNP5D* and *MNP6P* to be positive, but the former should be larger than the latter to the extent that MNP quality is important to consumers. *MNPR* is also expected to have a positive coefficient, but its success in explaining churn will depend upon how well its specific function form fits the data.

We also note that the decision to enact MNP regulation may be affected by market conditions, including churn levels. The econometric model will need to take this possible endogeneity into account.

The number of operators, *Ops*, should have a positive coefficient reflecting increased switching options and promotional activity as the number of operators rises.

A proxy for real incomes, *RGDPPC*, is included to a capture possible reduction in disconnections as income rises. This is expected to have a negative coefficient, because we expect that, in line with previous research, customers' demand for mobile network access is positively related to income. If this is the case, it also seems likely that users with higher income are *ceteris paribus* less likely to stop using mobile telephony once they have started than those of lower income. Since the churn figures include those who disconnect from one network without connecting to another, it is likely to be lower in markets with higher average incomes.

Finally, we include cellular density terms, which measure the number of mobile connections per head of population (*DEN).* This is intended to allow for a possible relationship between market maturity and churn. We might expect an increase in switching propensity as customers become more familiar with mobile telephony and as cohorts with greater price sensitivity take up access. In more mature markets, falling demand growth may weaken the incentives for switching by changing the nature of

competition (for example, via reductions in handset subsidies). Since density tends to approach a limit as each market matures rather than continuing to rise linearly, we include higher order transformations of this variable in the regressions.[110]

Quarterly dummies (*Q1* and *Q3*-*Q4*) are also included. Other factors that could affect churn, but on which data are not available, include the rate of service innovation, the extent of pre-paid customer registration vs. anonymity, the frequency of customer repeat purchase or sampling, contract lengths, the level of other (non-number-related) switching costs and the extent of substitutability between services of different operators.

We allow for I.I.D. errors in the measurement of variables through a disturbance term (ε_{it}). It also seems likely that data limitations, particularly regarding local preferences and service characteristics, have led to omission of variables that might help explain the level of churn in each country, so we expect to observe significant individual effects at country level (u_i).

Hence, for country i = 1...38 and quarter t = 1…20:

$$\ln\left(\frac{CHURN_{it}}{1-CHURN_{it}}\right) = \alpha + \beta_1 OPS + \beta_2 GDPPC_{it} + \beta_3 DEN_{i(t-1)} + \beta_4 DEN_{i(t-1)}^{\ 2} + \beta_5 DEN_{i(t-1)}^{\ 3} + \beta_6 Q1_{it} + \beta_7 Q3_{it} + \beta_8 Q4_{it} + u_i + \varepsilon_{it} + [\beta_9 MNP5D_{it} + \beta_{10} MNP6P_{it} \text{ or } \beta_9 MNPR_{it}] \quad (2)$$

Summary of prior expectations about coefficients:

$\beta_1, \beta_3, \beta_5, \beta_9, \beta_{10} > 0;$
$\beta_2, \beta_4 < 0$

[110] For a recent survey of empirical work on mobile telephony density, see Banerjee and Ros (2004).

Econometric Results

In this section, we estimate the model described in Equation 2 above. The results are shown in Table 15 below.

Table 15: Churn regression results using Arellano-Bond estimator, with MNP variables treated as endogenous

Variables and statistics	**Using MNP delivery time threshold dummies (<=5 days, >6 days)**		**Using reciprocal of MNP delivery time for countries with MNP**	
Dep. variable	$LGTCHURN_{it}$		$LGTCHURN_{it}$	
	Coef.	*Robust t-stat.*	Coef.	*Robust t-stat.*
$LGTCHURN_{i(t-1)}$	0.682	*12.12***	0.675	*12.03****
***MNP5d*$_{it}$**	**0.166**	***2.09****		
***MNP6p*$_{it}$**	**-0.171**	***-1.63***		
***RMNP*$_{it}$**			**0.00752**	***1.49***
OPS_{it}	-0.00719	*-0.17*	-0.0117	*-0.27*
$LRGDPPC_{it}$	-0.171	*-1.75**	-0.164	*-1.57*
$DEN_{i(t-1)}$	1.22	*1.5*	1.31	*1.95**
$DEN_{i(t-1)}^2$	-2.13	*-1.65**	-2.14	*-2.27***
$DEN_{i(t-1)}^3$	1.43	*1.9*	1.39	*2.64****
Constant	-0.00646	*-0.98*	-0.00623	*-1.02*
$Q1_{it}$	0.0227	*1.49*	0.0230	*1.46*
$Q3_{it}$	0.0210	*1.17*	0.0229	*1.27*
$Q4_{it}$	0.0409	*2.49***	0.0420	*2.54***
Sample	38 countries		38 countries	
Observations	667		667	
Min. periods	7		7	
Avg. periods	17.6		17.6	
Max. periods	20		20	
F(12,654)	35.9			
F(11,655)			32.4	
Arellano-Bond residual serial correlation test, order 2	Z = 0.04 [0.972]		Z = 0.03 [0.978]	

*Note: All variables are in first differences apart from the constant, and variables with an L prefix are in log terms. Figures in italics are t-statistics; *, ** and *** denote significant at the 10%, 5% and 1% level respectively. Numbers in brackets are p-values. Data sources: see Table 13 above.*

Since diagnostic tests after fixed effects OLS estimation showed evidence of autocorrelation and heteroscedasticity,[111] we estimated the models using the Arellano-Bond "difference GMM" estimator with robust standard errors. T-statistics are reported rather than Z-statistics due to the relatively small sample.

The one-period lag of our transformed churn variable is highly significant, positive and less than one, showing substantial persistence in the churn process. We find no evidence of second order autocorrelation in the residuals.[112]

There is a significant difference between the churn dummies for countries with a five day or shorter maximum target porting time and those permitting a longer porting time.[113] Countries requiring faster porting times experienced significantly higher churn rates after MNP, whereas there was no significant effect for those with a slower standard. Our alternative MNP variable based on the reciprocal of the target maximum porting time seems to have little explanatory power.

It is difficult to directly interpret the levels of coefficients in a model where the dependent variable has undergone a logistic transformation. However, in Table 16 below, we provide simulation results for the average treatment effect of MNP on quarterly churn rates and the equivalent increase in the average level of churn for countries with porting times of 5 days or less.

---

[111] Modified Wald test for groupwise heteroscedasticity: $\chi^2(38)=10,300$ [0.000]; Wooldridge test for autocorrelation in panel data: $F(1,37) = 17.3$ [0.0002]

[112] Second order autocorrelation would have been indicative of inconsistency, as per Arellano and Bond (2001), pp.281-282.

[113] A Wald test rejected equality between the MNP coefficients: $F(1,654) = 6.88$ [0.0098]

Table 16: Estimated MNP average treatment effect on churn and equivalent change in quarterly churn rates for countries with <=5 day porting rate target		
Measure	**Short-run**	**Long-run**
Average treatment effect[114]	0.253%	0.714%
Implied percentage change compared to sample average churn rate (2.05% as per Table 13)	+13.6%	+34.7%

The short-run predicted increase in churn seems consistent with predictions in *ex ante* studies. For example, in a CBA conducted for Hong Kong, scenarios were examined allowing for increases of 5-15% in the churn rate following introduction of MNP.[115] Our estimate is slightly lower than the 15% increase in actual switching (as opposed to churn) after MNP reported for South Korea in Kim (2005),[116] but we should expect this given that not all churn involves an inter-operator switch. However, note that because of the strong persistence we find in churn rates, the model predicts that the long-run impact of MNP on churn will be significantly higher.

Other results from the models are broadly as expected. Two of the quarterly dummies are not significant, but the Q4 dummy provides evidence of higher churn in the fourth quarter. This may reflect seasonal marketing activity or shifts in demand. The number of operators and the constant term were also found to be insignificant.

All other coefficients in the two models have the expected signs, although income is of only marginal significance. While the cellular density terms in the first model appear to

[114] The treatment effects are evaluated with other variables set to their sample averages. For the long-run effect, the current period and lagged churn rates converged at 1.77% with MNP, in comparison to 1.06% without MNP.

[115] NERA/Smith (1998), p.66.

[116] Non-switching status fell from 91% to 79.1% of those surveyed; Kim (2005), Table 2.

be individually insignificant, this is probably due to multicollinearity; a joint test on them rejects a zero value.[117]

We also tested lags of the MNP variables from 1-4 quarters, but the highest significance level was achieved with no lag.

5.5 Modelling the effect of MNP on prices

The cross-country data available for estimating the effect of MNP on retail prices limits us to a relatively simple modelling strategy. In particular, it is not possible to maintain the standard access/usage distinction and other more complex features of telephony demand models. Again we employ two models using different proxies for quality-adjusted MNP. These models are described below.

Shew (1994, p.51) set out a useful classification scheme for the determinants of a mobile operator's optimal price schedule. In his model, the minimum price P^* of a service bundle Q can be expressed as

$$P^* = f(\mathbf{S}, \mathbf{M}, \mathbf{R}, \mathbf{Q}) \qquad (3)$$

where[118]

- **S** includes a range of service quality characteristics such as area covered, probability of a call being blocked due to congestion and the firm's reputation. One might also include characteristics such as population coverage and probability of dropped calls.

[117] F(3,654) = 2.85 [0.0366]

[118] P^* is a minimum price in the sense that a given operator may offer several service packages with different price schedules, but for any given Q there is an optimum P^* among these packages.

- **M** includes market environment variables. Data on local cost conditions may be of importance to allow for possible economies of scale and density. Personal income levels, and measures of the extent of competition in each market and the availability of substitutes are further potential environmental factors.

  Other environmental factors flagged by Shew include commuting time, the maturity (or "age") of the market and the mix of industries in each area. He also points out that cost of living may affect both supply (through higher input costs) and demand for mobile telephony services.

- **R** is a vector of regulatory variables; in our analysis this would include the presence of number portability.

- **Q** includes characteristics of the service bundle, such as the number of minutes of calls included in it.

The price variable

The proxy for prices is quarterly real average revenue per minute (*RPM*). It is an aggregate measure encompassing all revenues associated with mobile voice services in each country (but excluding revenue from data services).

Use of an average revenue proxy for prices involves a departure from the approach used by Shew and most other analyses of regulatory impact in the mobile sector, such as Hausman (1995 and 2004). Shew's price schedule P^* was defined over **Q** at the level of each individual operator, and his empirical work compared prices set by operators for three given bundles of minutes.

RPM has some advantages as a price proxy. For example, we have already noted that charges for service components such as handsets and call termination may be affected by

MNP, and these might not be captured if we were to focus on some other measure, such as the average price of a three minute call or the price of a bundle of X minutes.

However, the benefits of aggregation come at a price. In particular, previous research into telephony demand has highlighted differences in the determinants of demand for network access and network usage (i.e. calls). *RPM* aggregates these differences away. Other potentially important features of telephony pricing are also obscured by averaging, including handset subsidies, time of day effects, innovation in tariff structures (e.g. bundling schemes and pre-payment offerings) and the mix of different call types (e.g. national vs. international).

Explanatory variables

We use the same two alternative sets of regulatory variables as in the analysis of churn described above. The first model includes dummy variables based on target maximum porting times: one where *MNPTM* was 5 days or less (*MNP5D*) and one where it was 6 days or longer (*MNP6P*). The second model uses the reciprocal of *MNPTM*.

Table 17 below shows how the average of *RPM*, our proxy for price of mobile services, varies in the sub-samples with and without MNP. These statistics paint a surprising picture, inasmuch as MNP appears to increase prices.

Table 17: Relationship between average prices (real revenue per minute) and mobile number portability

Case	Sample mean real revenue per minute (USD)
No MNP	0.192
MNP delivery time <= 5 days	0.206
MNP delivery time 6+ days	0.233

Source: see Table 13 above.

However, these descriptive statistics may be misleading. To illustrate some of the potential factors with a bearing on the MNP-price relationship, we refer to the correlation matrix in Table 23, Annex 5.1 below. First, the negative relationship between *lrpm* and *time* shows that there was a declining trend in consumer prices across all countries during the period. Where MNP was implemented it tended to come later in the time series. The resulting timing effect will tend to bias the MNP averages downward. A similar downward bias may arise because there was a positive association of MNP with quantity of call minutes sold (*ltotmin*) and a negative relationship between quantity and price. In contrast, GDP (*lgdppc*) was positively associated with both MNP and prices, which could lead to an upward bias in the average. To fully isolate the effects of MNP from other variables requires regression analysis.

Unlike the MNP coefficients in the churn models, the coefficients on MNP variables in the price models are expected to be negative, reflecting stronger competition in markets with lower switching costs. As in the switching model above, we treat MNP as potentially endogenous.

Since each price observation in our dataset is associated with a unique *Q* (unlike the studies cited above, which fixed *Q*), we must control for the total quantity of minutes actually delivered by operators in a given country at price *RPM*. This quantity variable is designated *TOTMIN*, and it too is treated as potentially endogenous to allow for the simultaneous determination of quantities and prices in mobile markets.

Detailed information on service characteristics is not readily available on an internationally-comparable basis. However, since we have time-series cross-section data, characteristics that are jurisdiction-specific may be captured by the use of individual effects.

A richer dataset is available on market environment factors (**M**). We have a choice of proxies for the effects of market concentration on pricing, including the Herfindahl-Hirschman index (*HHI*), the one-firm concentration ratio (*CR1*) and the number of network operators (*OPS*). These are tested alternately in the regression since we do not wish to prejudge the nature of competition in the market. If greater concentration implies weaker competition in mobile telephony markets, *HHI* and *CR1* should have positive coefficients when each of them is included, and *OPS* should have a negative one.

Population density (*PDNST*), a proxy for local cost conditions, should have a negative coefficient reflecting economies of density. Real GDP per capita (*RGDPPC*), a proxy for income, might take a positive coefficient as per the reasoning in Shew (1994, p.114) that customers in high income areas will exhibit less price sensitivity, leading to higher prices in such areas. Both of these variables might have a non-linear relationship to average prices, so higher order terms are included in the regression.

We also include a time trend (*TIME*) to allow for time-varying unobserved effects and quarterly dummies to capture seasonal variations in pricing policies and demand patterns.

Other potentially relevant variables were unavailable for the relevant set of countries and periods, including details of marginal price schedules, prices of substitutes (e.g. fixed line services), differences in contract terms, quantities of spectrum allocated in each country, the extent of trans-national ownership or control of operators, availability and relative importance of pre-paid services, advertising expenditure, and regulatory variables other than MNP (e.g. requirements to offer wholesale roaming or access to service providers).

As in the churn model discussed earlier, we include a disturbance term (ε_{it}) and control for individual effects at country level (u_i). Logs are taken of continuous variables, including *RPM*.

To summarise, for country i = 1...37 and quarter t = 1…20:

$$\begin{aligned}\ln(RPM_{it}) = {} & \alpha + \beta_1 \ln(TOTMIN_{it}) + \beta_2 HHI_{it} + \\ & \beta_3 \ln(PDNST_{it}) + \beta_4 \ln(PDNST_{it})^2 + \beta_5 \ln(PDNST_{it})^3 + \\ & \beta_6 \ln(RGDPPC_{it}) + \beta_7 \ln(RGDPPC_{it})^2 + \\ & \beta_8 Q1_{it} + \beta_9 Q3_{it} + \beta_{10} Q4_{it} + \beta_{11} TIME_t + u_i + \varepsilon_{it} + \\ & [\beta_{12} MNP5D_{it} + \beta_{13} MNP6P_{it} \text{ or } \beta_{12} MNPR_{it}]\end{aligned} \tag{4}$$

Summary of prior expectations about coefficients:

$\beta_2, \beta_4, \beta_6, \beta_{12}, \beta_{13} > 0;$
$\beta_1, \beta_3, \beta_5, \beta_7, \beta_{11} < 0$

Econometric Results

In this section, we estimate the model described in Equation 4 above. Table 18 below sets out the regression results. As we found when modelling churn, initial estimation using OLS with fixed effects gave rise to heteroscedasticity and autocorrelation.[119] Here too, we estimated the models shown below using the Arellano-Bond estimator with robust standard errors, and diagnostic testing rejects the presence of second order serial correlation in the residuals. Due to differencing of the data, the fixed effects are eliminated and the differenced time trend yields a constant.

[119] Modified Wald test for groupwise heteroscedasticity: $\chi^2(38)$=12,800 [0.000]; Wooldridge test of autocorrelation in panel data: F(1,36) = 27.8 [0.000]

Table 18: Price regression results using Arellano-Bond estimator, with *LTOTMIN* and MNP variables treated as endogenous

Variables and statistics	**Using MNP delivery time threshold dummies (<=5 days, >6 days)**		**Using reciprocal of MNP delivery time for countries with MNP**	
Dep. variable	$LRPM_{it}$		$LRPM_{it}$	
	Coef.	*Robust t-stat.*	Coef.	*Robust t-stat.*
$LRPM_{i(t-1)}$	0.453	*4.92***	0.458	*4.83****
$\boldsymbol{MNP5d_{i(t-1)}}$	**-0.0658**	***-3.18****		
$\boldsymbol{MNP6p_{i(t-1)}}$	**-0.0258**	***-0.94***		
$\boldsymbol{RMNP_{i(t-1)}}$			**-0.00676**	***-3.41****
$LTOTMIN_{it}$	-0.343	*-5.46****	-0.332	*-5.35****
$LPDNST_{it}$	-1.30	*-0.25*	-1.32	*-0.27*
$LPDNST^2_{it}$	1.08	*0.77*	1.10	*0.82*
$LPDNST^3_{it}$	-0.133	*-1.15*	-0.132	*-1.17*
$LRGDPPC_{it}$	1.20	*2.37***	1.22	*2.52***
$LRGDPPC^2_{it}$	-0.0306	*-1.29*	-0.0330	*-1.45*
$LHHI_{it}$	-0.158	*-1.35*	-0.169	*-1.34*
Constant	0.00231	*0.9*	0.000556	*0.23*
$Q1_{it}$	-0.0250	*-3.28****	-0.0263	*-3.45****
$Q3_{it}$	0.0115	*1.5*	0.0113	*1.49*
$Q4_{it}$	0.00344	*0.53*	0.00305	*0.47*
Sample	37 countries		37 countries	
Observations	649		649	
Min. periods	4		4	
Avg. periods	17.5		17.5	
Max. periods	20		20	
F(14,634)	974			
F(13,635)			1,290	
Arellano-Bond residual serial correlation test, order 2	Z = -0.90 [0.366]		Z = -1.02 [0.307]	

*Note: All variables are in first differences apart from the constant, and variables with an L prefix are in log terms. Figures in italics are t-statistics; *, ** and *** denote significant at the 10%, 5% and 1% level respectively. Numbers in brackets are p-values. Data sources: see Table 13 above.*

We found evidence that MNP reduces retail prices, but only when its quality is high. For those countries with MNP delivery times of five days or less, the estimated short-run effect of implementing MNP was a fall in real average prices of about 6.6%, after a one quarter lag.[120] The estimated long-run reduction was significantly higher, at 12%. We

[120] We tested lags of between 0 and 4 quarters on the MNP variables, and the signs were the same in all cases, although statistical significance varied. A one quarter lag yielded the highest t-statistic for *MNP5d* and is thus reported here.

also found a negative MNP coefficient for countries with longer MNP delivery times, but it was not significantly different from zero. However, these results fall short of proving that a tighter MNP standard yielded a stronger price effect; we could not reject the hypothesis that the coefficients on the two MNP dummies were equal (F(1,634) = 2.70 [0.101]).

The alternative approach of including the reciprocal of each country's maximum time for MNP delivery also yielded a negative coefficient. This model implies a substantial price effect in countries with the tightest MNP delivery standards, but little effect elsewhere (see Table 19 below).

Table 19: Estimated effect of MNP on real average retail prices from 1/*MNPTime* model

MNP standard	Short-run	Long-run
2 hours	-8.11%	-14.97%
2 days	-0.34%	-0.62%
5 days	-0.14%	-0.25%
10 days	-0.07%	-0.12%
20 days	-0.03%	-0.06%

As expected, we found a robust inverse relationship between the number of minutes of traffic and real average prices. Income and population density variables also had the expected signs. Although t-tests on each of the population density terms suggested a lack of statistical significance, a joint test on all the terms strongly rejected a zero value: F(3,634) = 3.59 [0.0135]

Neither HHI (shown above) nor alternative proxies for market concentration (*CR1* and *OPS*) proved to be significant. We did find evidence of lower average prices in the first quarter of each year, perhaps reflecting the effect of temporary discounts on packages sold in the fourth quarter, but the constant term is not significantly different from zero.

5.6 Conclusions

Our central finding is that prices fell and churn increased in countries with a five day or better MNP delivery standard, as summarised in Table 20 below.

Table 20: Estimated effect of MNP on real average retail prices and churn rates, for countries with a <= 5 day target maximum porting time

MNP standard	Short-run	Long-run
Average prices (real revenue per minute)	-6.58%	-12.0%
Churn rate, quarterly average	+13.6%	+34.7%

The price result can be compared to the finding in Viard (forthcoming) that there was a 4.4% fall in prices after the introduction of toll-free number portability, which is a different but similar service.[121]

We found no significant effect of MNP on churn or average prices for countries that applied a less stringent target for maximum porting time. For jurisdictions requiring "high quality" MNP, our results are consistent with the presence of significant Type 1 and Type 2 benefits.

Areas for further research

The mobile market data currently available on a consistent basis over time and across countries has limitations when used for modelling the effects of MNP. First, our choice of a five day porting time threshold for examining MNP quality is dictated by the need to have an adequate sample of observations above and below the threshold. Future research into the effect of MNP will also benefit from the existence of additional time series data from jurisdictions where MNP has been implemented; many countries in our sample had only recently introduced these services.

[121] See Section 0.5 above.

Also, publication of harmonised cross-country data by supranational bodies such as CEPT, which published most of the MNP implementation and porting time data we used in this chapter (see Table 21 in Annex 5.1), should make it easier for future researchers to make inter-country comparisons.

Second, we have not been able to control for the varying price of MNP across countries. In some jurisdictions, MNP is free to the subscriber. In others, it can involve significant fees. For example, the system adopted in Singapore in 1997 permitted operators to levy monthly charges on users, but from August 2003 onwards only a one-time administrative fee was allowed.[122] There is also variation in the levels of one-off fees among those jurisdictions that permit them to be charged.[123] While charging could act as a deterrent to usage of MNP,[124] published information on such charges and on other aspects of MNP quality (for instance, whether or not it covers SMS messages) is scanty, and these dimensions are not explicitly addressed in our analysis.

One extension we plan to carry out using existing data is to examine that pattern of market share convergence after new entry in mobile markets. As we shall see in the next chapter, concentration in these markets is underpinned by government regulation, in the form of administrative entry barriers used to manage scarce radio spectrum. When entry is permitted, market shares tend to move towards symmetry. However, the speed and completeness of convergence to symmetry are important parameters in debates about appropriate regulation of these markets. Theory suggests that switching costs should account for some of this persistence in market shares.

[122] Infocomm Development Authority, Singapore, 2003.

[123] Ovum (2005), Section 3.5.

[124] Buehler, Dewenter and Haucap (2006), p.395.

We intend to explore this effect by using an extended version of the Chapter 5 dataset to estimate dynamic concentration models akin to the single country models introduced by Levy (1985), but in our case exploiting information on cross-country variation as well. This analysis should be of assistance to regulators considering whether markets subject to consumer switching costs are "sufficiently" competitive, or whether further regulatory measures are required.

Annex 5.1: Additional information on the dataset used in Chapters 5 and 6

DEN is each country's average number of mobile telephony users (including both post-paid and pre-paid customers) divided by the country's population. This variable may be subject to varying reporting practices in different jurisdictions. While it is easy to define and measure the number of post-paid subscribers, in most jurisdictions these represent a minority of mobile telephony users. The identities of the remainder, who use mobile telephony on a pre-paid basis, are often unknown to their network operators. As a result, network operators generally use a formula to estimate the number of active customers, typically treating a subscriber as active if his or her phone has been used within a set number of months. While we understand from Merrill Lynch (2004) that these formulae may vary across the sample, we have no details of the differences.

This caveat also affects the *CHURN* variable, which is a quarterly average of monthly actual and imputed[125] disconnections from networks as a proportion of the average number of users in each period.

Gross domestic product in real USD terms per capita (*RGDPPC*) was calculated for OECD countries based on local currency real GDP figures and GDP deflators from the OECD quarterly national accounts database. Exchange rates were taken from IMF International Financial Statistics. Figures for non-OECD countries are taken from the IMF World Economic Outlook database (September 2004), and are annual data, rather than quarterly. This treatment of GDP in non-OECD countries is not ideal, but as no quarterly

[125] Pre-paid users inactive for a specified period.

national accounts data were available for these countries it was unavoidable. In any event, the coefficients on GDP are not the focus of our analysis.

Table 21: MNP sample coverage, porting times and introduction dates

Country	Churn observations	RPM observations	MNP implemented ("-" if not implemented by 2Q04)	Target maximum porting time (days)	Main source for date of MNP implementation
Argentina	10	20	-	-	
Australia	20	20	3Q01	0.0833	TIO (2002, p.45)
Austria	20	20	-	-	ECC/CEPT (2005)
Belgium	16	13	3Q02	2	ECC/CEPT (2005)
Brazil	20	20	-	-	
Canada	20	20	-	-	Telegeography (2005b)
Chile	20	19	-	-	
Colombia	16	12	-	-	
Denmark	8	20	3Q01	5	ECC/CEPT (2005)
Egypt	7	8	-	-	
Finland	20	20	3Q03	5	ECC/CEPT (2005)
France	20	20	3Q03	30	ECC/CEPT (2005)
Germany	20	20	4Q02	6	ECC/CEPT (2005)
Greece	20	20	3Q03	1	ECC/CEPT (2005)
Hong Kong	20		1Q99	2	OFTA (2006)
Hungary	12	19	2Q04	14	ECC/CEPT (2005)
India	20	20	-	-	Telegeography (2005a)
Ireland	11	11	3Q03	0.0833	ECC/CEPT (2005)
Israel	20	18	-	-	
Italy	20	20	2Q02	5	ECC/CEPT (2005)
Japan	20	20	-	-	
Malaysia	20	20	-	-	
Mexico	20	20	-	-	
Netherlands	20	20	2Q99	10	ECC/CEPT (2005)
New Zealand	20	20	-	-	
Norway	14	20	4Q01	7	ECC/CEPT (2005)
Poland	20	9	-	-	ECC/CEPT (2005)
Portugal	20	20	1Q02	20	ECC/CEPT (2005)
Russia	17	17	-	-	
South Africa	20	20	-	-	
Spain	15	20	4Q00	4	ECC/CEPT (2005)
Sweden	20	20	3Q01	5	ECC/CEPT (2005)
Switzerland	20		1Q00	5	ECC/CEPT (2005)
Taiwan	12	15	-	-	
Thailand	20	4	-	-	
Turkey	13	8	-	-	
UK	20	20	1Q99	9	ECC/CEPT (2005)
US	16	17	4Q03	0.104	FCC (2004)
Venezuela		19	-	-	

Table 22: Correlation matrix of variables used in MNP churn regressions

	lgtchurn	***mnp***	***mnp5d***	***mnp6p***	***rmnptm***	***ops***	***lrgdppc***
lgtchurn	1						
mnp	0.0309	1					
mnp5d	0.0051	0.731	1				
mnp6p	0.0385	0.556	-0.162	1			
rmnptm	0.0221	0.320	0.422	-0.0517	1		
ops	0.245	0.0698	0.0616	0.0258	0.0669	1	
lrgdppc	-0.297	0.472	0.348	0.258	0.150	0.242	1
lrgdppc2	-0.277	0.485	0.358	0.265	0.154	0.250	0.998
den1	-0.264	0.565	0.436	0.286	0.137	0.153	0.759
den1sq	-0.213	0.568	0.444	0.281	0.114	0.133	0.637
den1cu	-0.174	0.534	0.421	0.259	0.0863	0.130	0.537
time	0.0997	0.284	0.24	0.119	0.160	0.105	0.0412
q1	0.0152	0.011	0.0064	0.008	0.0045	0.0093	-0.001
q3	-0.0142	-0.0287	-0.0162	-0.0217	-0.0121	-0.0153	-0.0032
q4	0.0131	-0.0044	-0.0027	-0.0031	0.004	0.0009	0.005
	lrgdppc2	***den1***	***den1sq***	***den1cu***	***time***	***q1***	***q3***
lrgdppc2	1						
den1	0.759	1					
den1sq	0.641	0.9671	1				
den1cu	0.540	0.9098	0.984	1			
time	0.051	0.4217	0.456	0.468	1		
q1	-0.0006	0.0321	0.0328	0.0326	0.0508	1	
q3	-0.0049	-0.0658	-0.0689	-0.0693	-0.149	-0.33	1
q4	0.0048	-0.026	-0.0279	-0.0282	-0.0496	-0.334	-0.33

Note: Data described in Table 13, churn regression sample. Variables with a 1 suffix are lagged one period.

Table 23: Correlation matrix of variables used in MNP price regressions

	lrpm	***mnp1***	***mnp5d1***	***mnp6p1***	***rmnptm1***	***ltotmin***	***lhhi***
lrpm	1						
mnp1	0.210	1					
mnp5d1	0.0907	0.676	1				
mnp6p1	0.189	0.649	-0.122	1			
rmnptm1	-0.0186	0.340	0.484	-0.0413	1		
ltotmin	-0.151	0.129	0.0494	0.122	0.0789	1	
lhhi	0.203	-0.092	0.0162	-0.141	-0.0306	-0.545	1
lpdnst	0.152	0.152	-0.0558	0.262	-0.257	0.223	-0.135
lpdnst2	0.141	0.181	-0.0447	0.290	-0.192	0.262	-0.203
lpdnst3	0.123	0.192	-0.0452	0.305	-0.155	0.278	-0.250
lrgdppc	0.515	0.393	0.26	0.261	0.135	0.224	-0.114
lrgdppc2	0.504	0.405	0.267	0.270	0.139	0.234	-0.117
time	-0.429	0.298	0.284	0.108	0.157	0.234	-0.141
q1	-0.0495	0.0309	0.03	0.0107	0.0222	0.0023	-0.0082
q3	0.068	-0.062	-0.0644	-0.017	-0.0455	-0.0311	0.0261
q4	0.0459	-0.0036	0.0075	-0.0125	0.0023	-0.0158	0.0106
	lpdnst	***lpdnst2***	***lpdnst3***	***lrgdppc***	***lrgdppc2***	***time***	***q1***
lpdnst	1						
lpdnst2	0.979	1					
lpdnst3	0.940	0.989	1				
lrgdppc	0.0754	0.111	0.126	1			
lrgdppc2	0.0874	0.126	0.142	0.998	1		
time	0.0419	0.0367	0.0325	-0.0506	-0.0421	1	
q1	0.0041	0.0042	0.0042	-0.0077	-0.0078	0.0445	1
q3	-0.0053	-0.0052	-0.0054	-0.0018	-0.0026	-0.132	-0.327
q4	-0.0008	-0.0015	-0.002	0.0062	0.0058	-0.0507	-0.335
	q3	***q4***					
q3	1						
q4	-0.327	1					

Note: Data described in Table 13, price regression sample. Variables with a 1 suffix are lagged one period.

6 Accounting for ARPU[126]

Average revenue per user (ARPU) is a key metric used by analysts to track mobile telecommunications firms and markets. Higher ARPU may be the result of higher prices, greater minutes of use or a combination of the two. Some national regulatory authorities (NRAs) have argued that higher ARPU is the result of lack of effective competition – an NRA or telecoms regulator may believe that one or more mobile network operators (MNO) possess 'significant market power' (SMP), for which the test is broadly equivalent to that of individual or collective dominance under competition law. An operator with SMP or dominance under the new EU electronic communications regulatory framework is one capable of acting to a significant extent independently of its rivals, customers and ultimately of consumers. Collective or joint dominance refers to the ability of two or more undertakings to tacitly agree their behaviour, even though none of the operators may be singly dominant. Absent suitable regulation, either type of dominance would be expected to result in higher prices, less output or lower quality of service compared with effective competition.

The Irish electronic communications regulator, ComReg, made the first ever finding of a collectively dominant position in relation to European telephony in December 2004.[127] In arriving at its conclusion that Irish mobile operators possessed SMP in the Irish mobile access and origination market, ComReg noted an apparent correlation between ARPU and market concentration.

---

[126] This chapter is substantially based upon McCloughan, P. and Lyons, S., "Accounting for ARPU: New Evidence from International Panel Data", *Telecommunications Policy*, forthcoming.

[127] See Commission for Communications Regulation, 2004b. ComReg's proposed decision was confirmed by the European Commission but was later annulled following appeal proceedings, apparently on foot of procedural issues; see Electronic Communications Appeal Panel, 2005.

> "There appears to be a correlation between high ARPUs and high market concentration. Ireland, Switzerland and Norway are the countries with the highest levels of concentration, and which are also the countries with the highest levels of ARPU."[128]

An element of the debate in Ireland concerns whether high ARPU is primarily due to higher usage of mobile telephony services or the result of market power – an issue that has been a source of contention among regulators and mobile operators in other jurisdictions.

The view that concentration and other structural features of the market, including the conditions of entry, are a strong indicator of market power was widespread from the work of Bain (1951) up to the early 1970s (see Martin, 1994, pp. 196-214, for a review). However, following the lead of Demsetz (1973), who suggested that observed profitability differences between sectors might relate to the effects of intra-industry variation in efficiency, more recent studies have emphasised the importance of industry-specific structural factors (following the study by Schmalensee, 1989). Today this is reflected in the way that economists contribute to competition and regulatory assignments, in which a case-by-case approach is emphasised and effects-based modelling is applied.

This chapter empirically examines the arguments by developing a multiple regression model of ARPU that controls for structural factors affecting demand for mobile services and underlying factors conditioning the level of competition, which may affect prices. Estimated and tested using a unique panel of mobile telephony market data, the analysis permits examination of the following issues, *inter alia*, which have proved contentious between MNOs and their regulators.

---

[128] Commission for Communications Regulation, 2004a, para. 4.28.

First, whether higher market concentration leads to higher ARPU (given the presence of high entry barriers due to rationing of mobile telephony spectrum through national licensing schemes).

Second, the extent, if any, to which income may influence ARPU by affecting the demand for mobile services. In particular, higher national income per head might increase ARPU as users in more prosperous markets purchase higher minutes of usage than those in less prosperous markets;

Third, whether demographic factors might have any effect on ARPU. For example, if mobiles are used most intensively by the young, countries with higher youth dependency ratios might have higher ARPU because of higher overall usage levels.

Fourth, whether requirements on some operators to offer mobile number portability (MNP) reduce ARPU by making it easier for customers to switch between MNOs.

While other researchers have examined aspects of competition in mobile markets, empirical studies have tended to focus on explaining either the prices of service baskets (i.e. specified bundles of access and usage) or the average revenue per minute of calls. ARPU, in contrast, has received a great deal of regulatory and commercial interest, but relatively little academic attention. To the best of our knowledge, this chapter is the first attempt to empirically evaluate the determinants of ARPU.

The structure of the chapter is as follows. The next section reviews ARPU as a proxy for the price of mobile telephony services. The econometric model is formulated in Section 6.2. Section 6.3 describes the data used to estimate the model and the results are presented in Section 6.4. Section 6.5 discusses the policy implications of the results and concludes the chapter.

6.1 ARPU as a proxy for the price of mobile services

It is difficult to find a simple metric for the price of mobile telephony. Mobile pricing normally involves complex non-linear 'tariff schedules', which generally include a choice among various bundles of services combining minutes of different call types and, more recently, data services such as text messaging. As well as the price of the bundle, there is often a per-minute or per-text charge for services in excess of the bundle limits in a given period.

Empirical studies involving mobile prices normally attempt to collapse some of this complexity into simpler measures. One approach is to select one or more specific bundles of services and calculate the charge that would arise under each operator's tariff schedules.[129] For example, we could work out the bill that would arise under a given tariff package for 200 minutes of peak time national calls and 50 texts (or any other specified bundle). A variation on this approach that allows aggregation across all of an operator's tariff packages is to assume that each customer is on the 'optimal' package for his or her pattern of demand. This involves working out the locus of minimum charges across the bundles offered by an operator as quantity is varied (for a more detailed explanation, see Shew (1994, p. 50).

An alternative approach is to use average revenue as an indicator of the effective price. These measures represent the price of the actual quantities of service elements delivered in a given place and time. Unlike the basket-based measures, they are not linked to a specified quantity of services, which may be a disadvantage when it comes to modelling.

[129] See, for example, Hausman (2002), pp. 591-594, and Hausman (1995).

However, average revenue measures tend to include a wider set of service elements (e.g. termination revenue) than basket-based measures do.

ARPU, the focus of this chapter, is probably the average revenue measure most commonly cited by industry participants. Another measure is revenue per minute (RPM), which we have employed in Chapter 5 above. The difference between these measures is in the quantities that are being priced. RPM might be a better proxy for price if the purchase decision for those buying telephony services focused on the cost per minute. ARPU might be a more appropriate price proxy if mobile subscribers do not focus on the per-minute price, but instead compare the total cost of service offerings. Intuitive examples might be users that buy significant amounts of data services (which are not priced on a per-minute basis) or users that focus on their monthly bills rather than attempting to interpret mobile tariff schedules.

It is not immediately obvious which one of these approaches is the best one for summarising price when examining issues such as competition in mobile markets. In this chapter, we focus on ARPU in view of its widespread use in industry and regulatory circles. However, we will also refer to research using other price proxies where relevant.

6.2 Hypotheses and econometric model

To identify the main determinants of ARPU, we employ the framework described in Shew (1994) and discussed in Chapter 5 for classifying the determinants of mobile telephony prices.

Thus, the model states that the level of our price proxy ARPU in country *i* during quarter *t* *(*$ARPU_{it}$) is a function of vectors describing service quality, market environment, regulation and quantity. In more formal terms:

$$ARPU_{it} = g(\boldsymbol{S}_{it}, \boldsymbol{M}_{it}, \boldsymbol{R}_{it}, \boldsymbol{Q}_{it}) \quad (5)$$

The variables included in each class for the purposes of regression analysis are discussed below.

Service Quality Variables

Unfortunately, there is little direct internationally-comparable information on service quality factors at the quarterly frequency level employed in this study. However, by using a fixed effects estimator with longitudinal data, regression analysis of the type undertaken below should at least be able to control for the level of quality that is specific to a given country, but invariant over time. It seems unlikely that exclusion of time-varying quality factors materially affects the results.[130]

Market Environment Variables

Market Concentration

Traditional economic models of oligopolistic competition predict that more highly concentrated markets support higher supplier mark-ups and higher prices. For example, in the *n*-firm Cournot model, the market price-cost margin or Lerner index of market power is proportional to the Herfindahl-Hirschman index (HHI) of concentration.[131]

Empirical evidence on this relationship in mobile markets is mixed. In Chapter 5, we found that HHI and other concentration measures did not have a significant effect on mobile prices (proxied by RPM) when estimating a fixed effects model similar to the one described in this chapter. In contrast, Hazlett and Muñoz (2004), who estimated an

[130] Population coverage of mobile networks may indeed have a significant impact on revenue across *operators* (see Valletti & Cave (1998), p. 119, but it is unlikely to give rise to significant variations across *markets*, since at least one or two operators in each country normally have full coverage.

[131] See, for example, Shy (1995) for a treatment of oligopoly models.

inverse demand function for mobile telephony services to explore the welfare effects of spectrum allocation policies and also used RPM as a price proxy, found that HHI did have a significant impact on RPM.

In order to examine the relationship between ARPU and concentration as broadly as possible, three proxies for concentration are tested in separate regressions below.

The first regression takes the Herfindahl-Hirschman index to measure market concentration in country *i* during quarter *t* (HHI_{it}).[132] $CR1_{it}$ (the subscriber market share of the largest MNO) is used as a proxy for concentration in the second regression. This reflects the potential importance of first-movers in mobile markets, where the pattern of entry often reflects a single incumbent followed by a number of later entrants. A different approach is used in the third regression, which includes terms for the number of licensed operators: $OP2_{it,}$ $OP4_{it}$ and $OP5P_{it}$ (dummy variables for the number of network operators providing services, corresponding to 2, 4 and 5+ operators respectively).

If concentration is important in accounting for differences in ARPU across countries, as some national regulatory authorities have claimed, the estimated coefficients on HHI_{it} and $CR1_{it}$ should be positive. In modelling the effect of the number of operators, the dummy $OP3_{it}$ is treated as the base-case, so that the other dummies show the effects of a given number of operators relative to a three operator market. If concentration captured in this way is important, we might expect the coefficient on $OP2_{it}$ to be positive (i.e. ARPU in a two operator market is higher than that in a three operator market) and for the coefficients

[132] See footnote 131. The size variable is subscriber numbers.

on $OP4_{it}$ and $OP5_{it}$ to be negative, reflecting the traditional view that the ability to sustain a higher price is lower where the number of competitors is larger.

Population Density

It is well understood that mobile telephony costs may vary with the density of users on a network. Technology and regulatory constraints limit the maximum size of a mobile cell (the area served by a particular transmitter), while population density varies widely. To allow for the possibility that cellular networks exhibit economies of density, this study uses the log of population density ($LPDNST_{it}$) as a proxy for user density. However, in higher density areas such as cities, these economies might be exhausted. To allow for variation in the marginal effect of density on ARPU, the square of $LPDNST_{it}$ is also included.

If there are economies of density, $LPDNST_{it}$ is expected to take a negative coefficient in ARPU regressions (cost falls with density, leading to lower prices), while the squared term taking a positive coefficient may indicate density economies are exhausted beyond a high population density level.

Personal Incomes

The overall level of income in a country is a potentially important factor in determining the level of demand for mobile telephony services and hence ARPU. First, mobile telephony may be a normal or 'superior' good. Some past research suggests that the propensity to purchase mobile access is positively associated with personal incomes.[133]

[133] See, for example, Ahn and Lee (1999).

Second, the elasticity of demand may be different for operators serving higher income customers, leading to a different profit-maximising price for mobile operators as suggested by Shew (1994, p. 55). Moreover, the observed relationship between income and mobile usage may be affected by differences in the way services are used in developing and developed countries. For example, Banerjee and Ros (2004, p. 279) find that cellular density growth is faster in poor countries than in rich ones, and they attribute this to a tendency by less developed countries to build out mobile networks in preference to fixed networks.

The proxy used for income in this chapter is the log of real GDP per capita in US dollars ($LRGDPCAP_{it}$). If the relationship found in other research is repeated here, this variable should have a positive effect on $ARPU_{it}$. To allow for the possibility that the marginal effect of income on mobile telephony demand changes as income increases (e.g. because access reaches saturation), a squared term is also included. This latter term may be expected to have a negative effect on ARPU.

Prices of Substitutes or Complements

The prices of complementary or substitute goods may have a significant effect on demand. Fixed line telephony services are an obvious candidate for both of these categories: fixed line services share some characteristics with mobile services (and hence might be substitutable), but they are also linked by network effects (and hence might be complementary). Past studies have not reached a consensus on whether fixed line and mobile services are in general complements or substitutes. For example, Ahn and Lee (1999) find fixed and mobile network access to be complements, whereas Madden and Coble-Neal (2004) report strong substitution between fixed and mobile telephony using a dynamic model that allows for network effects in demand.

The proxies available for capturing the prices of fixed network services are less than ideal; only annual data (not quarterly) are available, and there are no published average revenue figures comparable to mobile ARPU or RPM. Nevertheless, two proxies are included here in an attempt to capture the effects of fixed services on mobile demand. For each country in the sample, the explanatory variable $RFIXRENT_{it}$ is the real average monthly rental charge in US dollars for fixed line services and $RFIXRATE_{it}$ is the real average price of a three minute fixed rate peak call. Both series are annual. Positive coefficients on either of these variables could indicate substitution between fixed and mobile services, whereas negative coefficients may be evidence of complementarity.

Churn

The churn variable, $CHURN_{it}$, reflects the proportion of subscribers that close (or stop using) accounts in a given period. This variable might have a negative effect on $ARPU_{it}$, insofar as higher churn is likely to be associated with higher market share volatility. This should reduce the scope for tacit collusion by making it harder for firms to distinguish between fluctuations in their sales caused by competitors' activities and variations caused by other (often unobservable) factors. Moreover, the set of customers closing or ceasing accounts is likely to be more open to competitive offers on average than customers in general. This would also imply that high churn should give rise to lower ARPU, other things being equal. On the other hand, there may be an argument that higher churn is associated with higher ARPU because higher churn leads to higher operator administration costs, which may have to be recouped through higher prices.

Demographic Effects

It is possible that some countries have significantly higher or lower mobile usage due to demographic influences on demand. The proxy used for this effect is the ratio of the number of people aged 15-24 to the total population: $YTHRAT_{it}$. To the extent that mobile

phone usage among younger age groups is higher than that of the general population, the sign on the explanatory variable $YTHRAT_{it}$ should be positive, other things being equal.

Time Effects

We also include a time trend ($TIME_{it}$) to allow for time-varying unobserved effects common to all countries (e.g. falling quality-adjusted costs of technology) and quarterly dummies $Q1_{it}$ - $Q4_{it}$ to capture any seasonal variations in pricing or marketing policies. $Q1_{it}$ is chosen as the base and is omitted in the regressions.

Regulatory Variables

Mobile Number Portability

The growing popularity and likely effects of this regulatory measure were discussed in detail in Chapter 5. Since MNP is not central to our analysis in this chapter, we simply include a dummy variable MNP_{it} indicating whether such a mandate has been imposed. Our prior is that MNP will have a negative relationship to ARPU, *ceteris paribus*.

Other regulatory measures that might affect ARPU, such as requirements to offer wholesale access (e.g. to mobile virtual network operators or resellers) or national roaming, proved impossible to incorporate in the analysis. This is partly due to limited availability of data on many countries, and partly because of the difficulty in defining complex forms of service provision on a consistent basis internationally.

Quantity Variables

In common with the approaches taken in Chapter 5 and Hazlett and Muñoz (2004), but differing from Shew (1994), the price proxy used in this chapter (ARPU) does not relate to a fixed quantity of services across all periods. The quantity supplied in a particular period (i.e. the number of users connected) is the outcome of a market process, so it will be necessary to control for this in the regression analysis.

Because the quantity of services (in this case, proxied by the number of users served) may both affect - and be affected by - prices, the number of mobile network users, $SUBS_{it}$, is taken as endogenous. This means that an instrumental variable is needed for $SUBS_{it}$ to order to derive reliable and consistent estimates. The instrument employed is a one period lag of the measure, or $SUBS_{i(t-1)}$. Because there tends to be strong persistence in the number of subscribers in a given market, this measure correlates closely to $SUBS_{it}$, with a correlation coefficient of 0.999 for the sample discussed in Section 6.3 below.

The log of $SUBS_{i(t-1)}$ ($LSUBS_{i(t-1)}$) is used in the regression analysis to facilitate interpretation of the coefficients. If there are significant economies of scale in the supply of mobile services, $LSUBS_{i(t-1)}$ should have a negative relationship to $ARPU_{it}$.

Econometric Model

Collecting the variables discussed above and taking the log of $ARPU_{it}$ as the dependent variable ($LARPU_{it}$) yields the following time-series cross-section regression model of ARPU in country i at time t:

$$\begin{aligned} LARPU_{it} = {} & \beta_0 + \beta_1 CONC_{it} + \beta_2 LSUBS_{i(t-1)} + \beta_3 LPDNST_{it} + \beta_4 LPDNST2_{it} + \beta_5 LRGDPPC_{it} \\ & + \beta_6 LRGDPPC2_{it} + \beta_7 MNP_{it} + \beta_8 CHURN_{it} + \beta_9 YTHRAT_{it} \\ & + \beta_{10} RFIXRENT_{it} + \beta_{11} RFIXRATE_{it} + \beta_{12} TIME_{it} + \\ & \beta_{13} Q2_{it} + \beta_{14} Q3_{it} + \beta_{15} Q4_{it} + u_i + e_{it} \end{aligned} \qquad (6)$$

The sample includes n countries over T quarters. Included is a country fixed effect term u_i, and the error term is specified as a classical disturbance – i.e. $e_{it} \sim IIN(0, \sigma^2) \;\; \forall \; i = 1,\ldots,n$ and $t = 1,\ldots,T$.

6.3 Data used for modelling

To estimate and test the econometric model specified in Equation 6 above, we once more employ an extended version of the dataset used in Chapter 5. Table 24 below provides

descriptions, summary statistics and data sources for the variables used in the econometric analysis.

Table 24: ARPU model variable descriptions, sources and summary statistics (balanced panel of 14 countries, 21 quarters; 294 observations)

Variable	Description	Source	Mean	St Dev	Min.	Max.
$ARPU_{it}$	Real monthly service revenues (US$) divided by average monthly mobile network user base	ML, deflated as per Annex 5.1	33.1	13.88	2	84
$SUBS_{it}$	Number of mobile telephony users (000s)	ML	14700	20,400	907	87,700
$PDNST_{it}$	Population density: population per Km^2	See Annex 5.1	1,030	2,220	2.47	6,970
$RGDPPC_{it}$	Real GDP per capita	See Annex 5.1	16,400	10,900	442	40,200
$CHURN_{it}$	Number of users disconnected from each network expressed as % of MNO's average number of users in quarter *t*. Average in *i*	ML	0.0235	0.0129	0.00868	0.0756
HHI_{it}	Herfindahl-Hirschman Index: Sum of the squares of the market shares (mobile network users) of all MNOs.	Own calculations based on ML	3,870	870	1,780	6,210
$CR1_{it}$	The top MNO's market share (of subscribers) in *i* at *t*.	Same as above	0.489	0.0943	0.253	0.746
$OP2_{it}$	=1 if two mobile network operators (MNOs) were active in *i* at *t*, zero otherwise	Same as above	0.122	0.328	0	1
$OP3_{it}$	=1 if three MNOs; zero otherwise (base)	Same as above	0.412	0.493	0	1
$OP4_{it}$	=1 if four MNOs; zero otherwise	Same as above	0.276	0.448	0	1
$OP5P_{it}$	=1 if five or more mobile MNOs; zero otherwise	Same as above	0.190	0.393	0	1
$YTHRAT_{it}$	% Population aged 15-24 ('youth ratio') each year in country *i*	US Census Bureau (2005)	0.147	0.0292	0.111	0.214
MNP_{it}	= 1 if mobile number portability	See Annex 5.1	0.269	0.444	0	1
$FIXRENT_{it}$	Real average residential fixed line rental each year (US$)	ITU (2004), deflated as per Annex 5.1	9.87	4.68	2.20	21.9
$FIXRATE_{it}$	Real average price of a three minute local fixed line peak call each year (US$)	Same as above	0.0739	0.0542	0	0.185

Notes: Merrill Lynch (2004) is referred to as ML.

As noted above, the Merrill Lynch data contains information on 46 countries. However, for many of these countries, the relevant series were incomplete. After linking the Merrill Lynch data with sources for the other variables of interest, a balanced panel of 21 quarters (from the second quarter of 1999 through the second quarter of 2004) was available for 14 countries, which are listed in Table 25 below. These countries are represented in the analysis that follows. Additional data information, including a discussion of the GDP deflators used to derive the real variables, is given in Annex 5.1 above.

Table 25: Countries represented in ARPU panel dataset

1. Australia	8. Japan
2. Austria	9. Malaysia
3. Finland	10. New Zealand
4. Germany	11. Portugal
5. Greece	12. Singapore
6. Hong Kong	13. South Africa
7. India	14. Thailand

6.4 Econometric results

Fixed effects OLS regression results are set out in Table 26 below. As discussed in Section 3.2 above, three variants of Equation 6 above were estimated. Each variant includes a different proxy for concentration: HHI_{it}, $CR1_{it}$ and the number of network operators ($OP2_{it}$, $OP4_{it}$ and $OP5P_{it}$).

As well as reporting the estimated coefficients for Equation 6, the goodness-of-fit statistics (R^2 and F-ratio values) and model diagnostics (tests for heteroscedasticity and residual autocorrelation) are also provided. The reported t-tests of hypotheses are based on robust standard errors accounting for apparent heteroscedasticity; autocorrelation is not a

problem. The goodness-of-fit measures indicate a model that is well-determined and explains ARPU well. In particular, the model explains almost 90% of the cross-country variation in $LARPU_{it}$. The estimated coefficients are similar across the three estimated equations, suggesting that the econometric model is robust.

Table 26: ARPU panel regression results – fixed effects OLS

Variables and statistics	**HHI as proxy for concentration**		**CR1 as proxy for concentration**		**Number of operators as proxy for concentration**	
Dep. variable	$LARPU_{it}$		$LARPU_{it}$		$LARPU_{it}$	
	Coef.	*t-stat.*	Coef.	*t-stat.*	Coef.	*t-stat.*
Constant	1.69	0.22	3.02	0.37	-0.0736	-0.01
$LHHI_{it}$	-0.183	-1.5				
$CR1_{it}$			-0.323	-1.37		
$OP2_{it}$					0.0516	0.82
$OP4_{it}$					-0.0104	-0.34
$OP5P_{it}$					0.0136	0.16
$LSUBS_{i(t-1)}$	-0.648	-14.02***	-0.633	-13.63***	-0.646	-9.61***
$LPDNST_{it}$	-3.533	-2.04**	-4.31	-2.63***	-4.39	-2.23**
$LPDNST2_{it}$	0.4571	3.76***	0.501	4.73***	0.573	4.64***
$LRGDPPC_{it}$	1.93	2.17**	1.85	2.05**	2.11	2.32**
$LRGDPPC2_{it}$	-0.0476	-1.07	-0.0436	-0.96	-0.0575	-1.22
$Q2_{it}$	-0.00303	-0.16	-0.00408	-0.22	-0.00488	-0.26
$Q3_{it}$	0.0195	0.96	0.0177	0.86	0.0190	0.93
$Q4_{it}$	-0.0132	-0.39	-0.0158	-0.46	-0.0132	-0.4
$YTHRAT_{it}$	-0.131	-0.04	0.532	0.18	1.97	0.76
$TIME_t$	0.00652	1.88*	0.00669	1.96*	0.00904	2.69***
MNP_{it}	-0.0379	-0.96	-0.0350	-0.88	-0.0372	-0.83
$CHURN_{it}$	-2.37	-0.72	-2.33	-0.72	-2.32	-0.79
$RFIXRENT_{it}$	-0.00929	-0.75	-0.00929	-0.76	-0.0128	-1.25
$RFIXRATE_{it}$	-1.76	-2.1**	-1.71	-2.06**	-1.52	-1.67*
Sample	14 countries		14 countries		14 countries	
Periods	21		21		21	
Observations	294		294		294	
Adj. R^2	0.899		0.899		0.898	
F-test (15,265)	91.2 [0.000]					
F-test (15,265)			87.2 [0.000]			
F-test (17,263)					78.5 [0.000]	
Heteroscedast-icity: Wald χ^2 (14)	3,950 [0.000]		4,030 [0.000]		3,840 [0.000]	
Autocorrelation F(1,13)	1.22 [0.290]		1.23 [0.288]		1.80 [0.203]	

*Note: t-statistics are robust, based on the Huber/White/sandwich estimator of variance. *, ** and *** denote significant at the 10%, 5% and 1% level respectively. Numbers in brackets are p-values. Heteroscedasticity test is a modified Wald test for groupwise heteroscedasticity; see Greene (2000, p. 598). Autocorrelation test is a Wooldridge test for autocorrelation in panel data; see Wooldridge (2002, pp. 282-83). Data sources: see Table 24 above.*

Market concentration has no statistically significant effect on ARPU, whether concentration is measured by the HHI, one-firm concentration ratio or the number of MNOs. Taken together with studies of other mobile price proxies such the model of RPM in Chapter 5, this suggests that concentration is not an important determinant of prices (as proxied by average revenues) in national mobile telecoms markets.

Four variables appear as important determinants of ARPU from our analysis. First, higher number of mobile telephony users in a market lowers ARPU. To interpret this finding, note first that the effect arises from the total number of users, not a market-specific measure such as the rate of penetration. This probably rules out potential demand side explanations, such as the possibility that marginal usage of telephony services declines as the market becomes saturated. Instead, the result might be explained by economies of scale, i.e. declining marginal costs as mobile networks grow. The results imply that a 1% increase in the (lagged) number of subscribers is associated with a 0.63-0.65% decrease in ARPU.

Second, increased population density also lowers ARPU for most of the range of population density, which supports the view of economies of density. However, at very high levels of population density, the relationship between population density and ARPU is positive ($LPDNST2_{it}$ is positive and significant in each of the three estimated equations). However, the latter effect accounts for very few observations.

Third, higher real GDP per capita increases ARPU (in a linear fashion). According to the estimates, every 1% increase in real GDP per capita leads, on average, to a rise in ARPU of 2%. This finding suggests that mobile telephony services are a "superior" good.

Fourth, there is evidence that the price of fixed telephony services is negatively associated with mobile ARPU. In particular, a 1% increase in the average price of a three minute

local fixed line call is associated with a 1.5-1.8% reduction in ARPU, other things being equal. This finding of a negative coefficient on $RFIXRATE_{it}$ suggests that fixed and mobile services may be complements (at least within the countries and quarters in our sample).

Other variables, including MNP, fixed line rental prices, the youth ratio and churn do not emerge as significant. The coefficients on the time variable ($TIME_{it}$) are positive and at least marginally significant in the three estimated equations; this implies that the average level of ARPU was increasing over the period considered.

6.5 Policy implications and conclusions

It is natural for regulators, for which analytical resources are costly, to use the simplest reliable indicators when assessing technical questions such as when a market is effectively competitive. However, research into market structure and performance since the 1970s has shown that there is often no reliable substitute for constructing a case-specific model capturing structural and other features of a specific market. In particular, the straightforward relationship between concentration and market power posited by early studies has not fared as well in more recent research.

This chapter has examined a wide range of factors thought to influence ARPU, which is used as a key metric in assessing price and performance in mobile telephony markets by national regulatory authorities and others (including investors). This chapter has included as explanatory variables the extent of concentration in national mobile markets and structural demand-side variables such as national income. Almost all of the national mobile telephony markets in this sample would be deemed 'highly' concentrated

according to the HHI threshold of 1,800 used internationally by antitrust authorities.[134] However, no evidence was found that market concentration, whether measured by the HHI, top firm concentration ratio or by the number of firms, has any influence on ARPU.

In contrast, the results indicate that income (proxied by GDP per capita) is likely to have a significantly positive effect on ARPU, while population density appears to be negatively related to ARPU at a declining rate.

On further reflection, the finding that concentration may not be systematically related to prices in mobile telephony markets is not that surprising. Unlike non-networked markets, where concentration may bear a simple relationship to the extent of competition, the market structure (i.e. number of MNOs) in mobile telephony markets is fixed externally – for example by spectrum licensing. Indeed, as Hazlett and Muñoz (2004, p. 11) note, the magnitude of concentration (HHI) is "largely the product of regulatory design." In the long-run this is obviously true, but it may be also the case in the short-run as incumbent operators tend to lose share, and entrants gain, as the market tends toward equilibrium.

Concentration in these markets may be more an artefact of regulation and timing than an indication of sustainable market power. Moreover, because of the network property of mobile markets, in which greater consumer utility is afforded through more subscribers on a given network, competition may still be intense in mobile markets (through price and non-price means), even though concentration is 'high' according to measures such as the HHI.

---

[134] See the EC Merger Regulations, European Council (2004), for a recent example.

The absence of a role for concentration in determining ARPU (or RPM, as noted in Chapter 5 above) may imply that retail mobile telephony markets are not prone to oligopolistic pricing behaviour. This position is consistent with that taken by some regulators, but not all of them. For instance, in 2003, the UK regulator Oftel (now Ofcom) concluded that the UK mobile market was not characterised by collective dominance, even though the market was highly concentrated. However, in December 2004, the Irish regulator, ComReg, provisionally deemed that the Irish mobile access and origination market is characterised by collective dominance. This conclusion was partly based on ComReg's analysis of the relation between ARPU and the HHI in a sample of EU countries. However, ComReg did not undertake any regression analysis to ascertain whether the relation was statistically significant. The evidence presented here suggests that there is no significant relationship between them, after controlling for various other explanatory variables.

Regulatory policy regarding mobile telephony markets, and other electronic communications sectors, is developing rapidly, in the EU and elsewhere. The analysis of this chapter, which draws together policy debates currently underway in the EU and academic modelling, suggests that regulatory authorities should undertake in-depth economic and econometric analysis as a foundation to their decisions in assessing the level of competition in mobile and other electronic communications markets. To the best of our knowledge, few NRAs in the EU are conducting detailed economic modelling but instead tend to rely on simple forms of benchmarking. The requirement to undertake well-founded analysis is all the more important given the regulatory obligations that can follow on foot of a finding of SMP.

7 Small Firms' Perceived Constraints as an Indicator of Relative Regulatory Burdens across Europe

This chapter asks whether survey evidence on perceived constraints can help explain differences in new small firms' survival rates across countries, firm sizes and broad sectors within Europe. To do this, we construct a new dataset combining European Network for Social and Economic Research (ENSR) survey evidence on the perceived constraints on small firms with Eurostat data on firm demography in ten European countries.

We find that some reported constraints help to explain variations in post-entry performance of new small firms, with (as expected) groups of firms that report constraints exhibiting lower survival rates. However, the relationship seems to be relatively weak and it is statistically significant for only a subset of perceived constraints (e.g. "Lack of skilled labour" and "Other"). Surprisingly, we find no significant effects of administrative regulation and access to finance constraints on survival rates, even over a three year period.

Until more convincing evidence is found showing that this type of survey evidence helps explain market outcomes, using it in the assessment or design of public policy may be problematical. In our conclusions, we suggest refinements in survey design that might improve the usefulness of perceived constraints data for use in making inter-country comparisons of small business burdens.

7.1 Survey evidence as an indicator of small business burdens

Most governments try to favour small businesses in various ways. Common policy measures include VAT registration thresholds and exemption from regulations based on firm size (e.g. exemption from audit requirements). Asymmetrical treatment of this kind may be justified by reference to market failures and asymmetries in the impact of regulation, taxation or other policies, but ultimately the question of whether welfare will

be improved by giving small businesses different treatment from larger firms must be an empirical one.[135]

One frequently-used approach to identifying (and sometimes quantifying) the extent of burdens on small firms is to ask the affected parties. This involves including questions about perceived burdens in surveys of small businesses. Examples of European studies of this kind are EIM/ENSR (1995) and Kox (2005), which extrapolated estimates made of the costs of administrative burdens in the Netherlands to a pan-European basis.

Kingston University (2005) identified many studies employing this type of evidence and expressed concerns about a general failure to link perceived burdens with outcomes:

> "The major problem with this kind of survey data is that it only scratches the surface as to *how* regulation might generate adverse consequences for small business owners. It provides little evidence of the processes through which regulation has its effects, good and bad, on small businesses or *why* business owners are satisfied or dissatisfied with regulations."[136]

In principle, it is possible to test how far perceived constraints, as reported in such surveys, represent actual impediments to some aspect of small business performance, although few studies seem to have attempted this. One study that did is Bartlett and Bukvič (2001), a single-country study focusing on Slovenia. Estimating models of SME employment growth, they found significant negative effects on growth associated with indices of perceived bureaucracy, cost of credit and labour taxes. However, being a single-country study, this paper is open to the possible criticism that in the absence of significant cross-sample variation in the institutional environment, the institutional effects detected may have more to do with unobserved differences in firms' characteristics than real differences in barriers.

There seems to be little past research into the explanatory power of perceived small firm constraints across a range of countries, despite the apparent benefit of having additional variation in policies across national borders. Perhaps the nearest examples are studies,

---

[135] Storey (1994), pp.254-257.

[136] Kingston University (2005), p.8.

often using data from regulatory "scorecards" or indices,[137] that test the effects of regulatory institutions on a range of market performance variables, e.g. Freeman (2002), Hemmings *et al.* (2002), Brandt (2004b) and Loayza (2004).

This chapter is related to empirical work both in the institutional theory tradition and in industrial organisation (and in particular, studies of the determinants of new firm survival).

A few recent studies have considered the interface between firms' perceptions and institutional arrangements. Pierre and Scarpetta (2006) considered how far measured levels of regulation are reflected in firms' reported perceptions, and Brunetti *et al.* (1998) examined how firms' reported perceptions of obstacles to doing business vary by country and region. Aidis (2005) looked at correlations among different types of institutional barriers perceived by firms in Lithuania, examining in particular the interactions between formal and informal barriers to growth.

There is, of course, a more extensive empirical literature concerning the relationship between structural features of industries and post-entry performance, at a level that tends to abstract from institutional arrangements (focusing instead on features such as presence of economies of scale, innovative behaviour, demand growth, or small firms' endowments of human or managerial capital). Contributions to this literature include Audretsch (1991, 1995), Audretsch and Mahmood (1994, 1995), Mata and Portugal (1994, 2002), Boeri and Bellman (1995); Mata *et al.* (1995), Storey and Wynarczyk (1996), Agarwal (1998), Harhoff *et al.* (1998), McCloughan and Stone (1998), Audretsch, *et al.* (1999), Honjo (2000), Mahmood (2000), Tveterås and Eide (2000), KPMG *et al.* (2002), Bartelsman *et al.* (2003) and Persson (2004). One institutional feature that has received some attention is availability of financing (e.g. Åstebro and Bernhardt (2003)).

In the remainder of the chapter, we describe and test a set of hypotheses about the relationships between perceived constraints on small firms and their survival prospects, controlling for other influences on survival.

[137] For two examples, see OECD (2001b) and the World Bank Doing Business database.

7.2 Model description

This section discusses the data and empirical models used in the chapter.

Data

To examine the effects of perceived business constraints on demographic outcomes, it was necessary to combine two unrelated[138] datasets. Evidence on perceived constraints was drawn from the 1999, 2001 and 2002 surveys conducted by the European Network for Social and Economic Research (ENSR) in the framework of the Observatory of European SMEs.

Here is the wording of the business constraint question posed by the ENSR surveys, together with the variable name we will use to refer to each answer (in italics):

> "Which of the following factors has been the major constraint on your business performance over the last two years?
>
> (READ OUT; ONLY ONE ANSWER ALLOWED)

labour	Lack of skilled labour
finance	Access to finance
newtech	Implementing new technology
organisation	Implementing new forms of organisation
quality	Quality management
regulation	Administrative regulations (on environment, health and safety etc.)
infrastruct	Infrastructure (road, gas, electricity, communication, etc.)
other	(DO NOT READ) Other. [NB: for the analysis in this chapter, we have included in this category two other responses: 'Introduction of the EURO' in the 2001 and 2002 surveys and 'Purchasing power of customers' in the 2003 survey.]
noburden	(DO NOT READ) None at all"

[138] In terms of design and implementation.

Some descriptive statistics from the 2003 survey are shown in Annex 7.1 below.[139] These surveys were sponsored by the European Commission, and the data are held by EIM in the Netherlands.[140] Observations were drawn from SMEs (enterprises with less than 250 employees) in 19 European countries. Each ENSR survey was stratified, with a target of about 8,000 total observations. The actual sample sizes varied; e.g. the 2001 survey covered 7,662 enterprises.

Business demography data, including information on enterprise birth and survival rates, is published by Eurostat as part of its "Business demography statistics" series. These data are based primarily on analysis of business registers by national statistical agencies, and according to Eurostat they are designed to capture "real enterprise births (and deaths), that is, enterprise births (deaths) that amount to the creation (dissolution) of a combination of production factors and where no other enterprises are involved. In other words, enterprises created or closed solely as a result of e.g. restructuring, merger or break-up are not included in this data."[141] Coverage varies by time and country, but the best samples of firm survival data relate to 2000-2002, and these are used in this chapter.

Although the variables in these two datasets were in most cases defined differently, it proved possible to identify a set of countries, time periods, firm size bands and industry sectors to which both datasets could be aggregated. The variables used in the analysis are listed in Table 27 below, including sources where relevant.

[139] See KPMG *et al.* (2004a) for more details of the survey.

[140] We are grateful to the European Commission and EIM for providing access to these data.

[141] Eurostat metadata on the dataset: http://europa.eu.int/estatref/info/sdds/en/sbs/bus_demo_base.htm. See Brandt (2004a) for a more detailed discussion of this dataset.

Table 27: Variables included in the combined business constraints dataset

Survival rates for firms of each type (Eurostat)	
surv1	One-year survival rate: Firms surviving since birth in year t-1 divided by total firms born in year t-1
surv2	Two-year survival rate: Firms surviving since birth in year t-2 divided by total firms born in year t-2
surv3	Three-year survival rate: Firms surviving since birth in year t-3 divided by total firms born in year t-3
Main constraint: % of firms of each type reporting the following as main constraint on business performance (ENSR 1999, 2001 and 2002).	
labour	Lack of skilled labour
finance	Access to finance
newtech	Implementing new technology
organisation	Implementing new forms of organisation
qualmgmt	Quality management
regulation	Administrative regulations (on environment, health and safety etc.)
infrastruct	Infrastructure (road, gas, electricity, communication, etc.)
other	Other (incl. 'introduction of the EURO' in 2001 and 2002 and 'purchasing power of customers' in 2003)
noburden	None at all (unprompted)
Birth rate of firms of each type (Eurostat)	
birthrt	Number of births of enterprises in year divided by number of enterprises active in year
Country dummies (1/0)	
dk	Denmark
es	Spain
fi	Finland
it	Italy
lu	Luxembourg
nl	Netherlands
no	Norway
pt	Portugal
se	Sweden
uk	United Kingdom
Year dummies (1/0)	
y2000	Year 2000 Eurostat and 1999 ENSR
y2001	Year 2001 Eurostat and ENSR
y2002	Year 2002 Eurostat and ENSR
Main activity dummies (1/0)	
manufact	ENSR: manufacturing industry; Eurostat: NACE D (manufacturing)
construction	ENSR: construction; Eurostat: NACE F (construction)
wholesale	ENSR: wholesale trade; Eurostat: NACE G51 (wholesale trade and commission trade, except of motor and motorcycles)
retail	ENSR: retail trade; Eurostat: NACE G52 (retail trade, except of motor vehicles, motorcycles; repair of personal and household goods)
hotels_cater	ENSR: hotels, catering; Eurostat: NACE H (hotels and restaurants)
transp_comms	ENSR: transport, communications; Eurostat: NACE I (transport, storage and communication)
finserv	ENSR: banking, finance and insurance; Eurostat: NACE J (financial intermediation); (1/0)
businserv	ENSR: business services; Eurostat: NACE 'K_not_K7415' (real estate renting and business activities excluding holding companies)
Size of firm dummies (1/0)	
empl_1_4	Firms with 1-4 employees
empl_5_9	Firms with 5-9 employees
empl_10p	Firms with 10 or more employees

The composite dataset covers 10 European countries, three years, eight sectors, and three firm size bands. Each combination of these dimensions defines a cell, hereafter referred to as a "firm type" and data on new firm survival rates, main constraints reported by firms, and firm birth rates represent cell mean values.

When combining the two datasets, judgement was required on how to match their various dimensions. Enterprises' countries of origin matched straightforwardly, and all countries that were represented in both datasets were included in the combined dataset. However, there were no observations of three-year survival rates for the Netherlands, Portugal and UK, so these countries were included only in the one- and two-year survival analyses.

The Eurostat demographic data cover 2000-2002, but matching ENSR surveys were available for the latter two years only. As a substitute for ENSR results from 2000, we used the 1999 survey.

While both source datasets have more disaggregated size bands, some aggregation was required to arrive at a common set. It includes three bands, covering firms with 1-4, 5-9 and 10+ staff. In cases where only partial coverage of a given band in either survey was available, we omitted the relevant cell from the analysis. In the ENSR results, these bands relate to the size of the firm in the preceding year, while in the Eurostat survey they refer to the current year.

Table 30 in Annex 7.1 provides descriptive statistics on all variables. Separate statistics are shown for three sets of observations, titled Samples A-C. These include all observations for which one-, two- and three-year survival rates were available, respectively, together with all explanatory variables.

Out of 720 possible cells,[142] missing data limited Sample A to 493 observations, Sample B to 364 observations and Sample C to 234 observations. Most of the missing cells arose from gaps in the Eurostat data.

[142] 720 cells = 10 countries X 3 years X 8 sectors X 3 firm size bands.

Modelling approach

The standard approach in empirical studies of new firm survival since the early 1990s has been to estimate survival probabilities or hazard rates using micro data on specified firm or market characteristics to explain survival patterns among individual enterprises (e.g. Audretsch and Mahmood (1994)). However, the focus of this chapter is on identifying the effects of institutional features that tend to vary more across national borders than within single jurisdictions, and only aggregate survival data are available on a pan-European basis.

Given the data limitations, a simple modelling strategy is employed. We aim to measure the average treatment effects of specified impediments to firms' likelihood of survival, using survey evidence on perceived performance constraints as a proxy for actual impediments, controlling for other factors suggested by the literature where data permit and relying on industry, country and year fixed effects to capture the remaining influences on post-entry survival. Our approach is thus more closely related to aggregate studies, e.g. Audretsch (1991) and Brandt (2004b), than to micro-data studies.

Dependent variable

The focus of this analysis is on survival of new firms. Questions about perceived constraints might also convey some useful information about other aspects of market performance such as entry rates. However, the questions in the ENSR survey were asked only of firms *in the market*, not potential entrants. We suspect such data are not well suited to explaining entry barriers; firms that did not enter are not represented in the surveys.

The survival variables for which reasonable numbers of observations are available from Eurostat are average one, two and three year survival rates for each firm type.

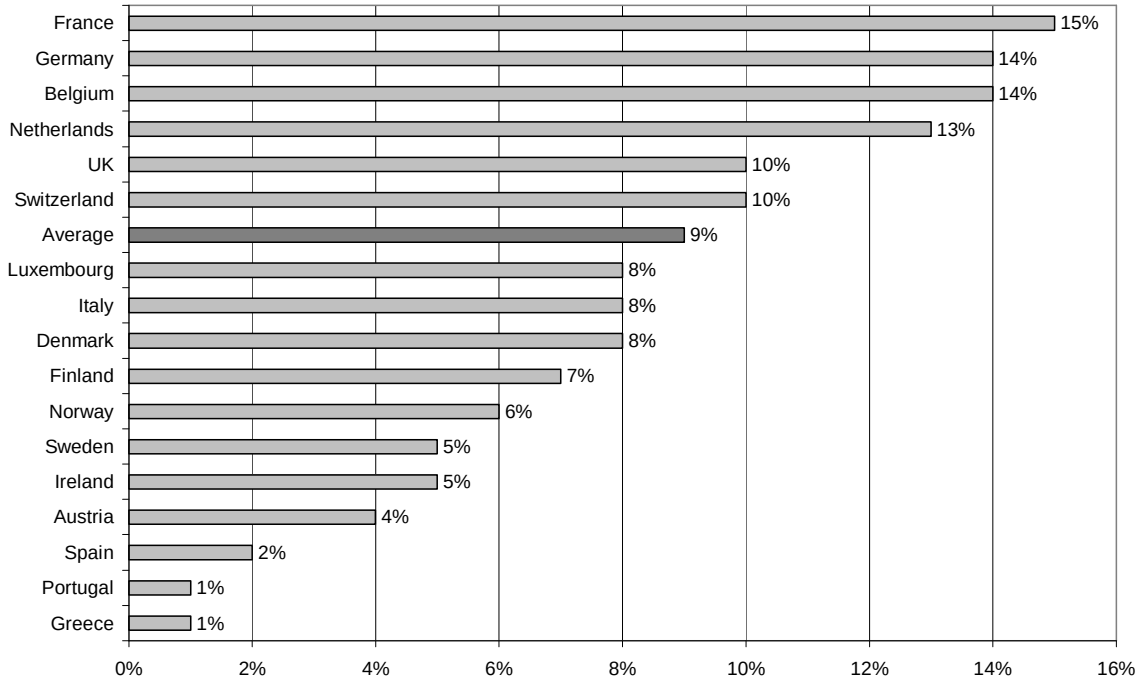
Explanatory variables

The main variables of interest are those capturing perceived constraints on businesses' performance. In particular, we expect to find that constraints associated with regulation and access to finance, which tend to elicit a policy response, are negatively associated with survival rates. Questions are included in the ENSR surveys on these and several other perceived constraints; the precise wording of these questions is provided earlier in this section.

The ENSR questions address relatively broad classes of constraints, and respondents were asked to identify the single major constraint faced, rather than scoring the relative importance of constraints in parallel with one another. Nevertheless, there is considerable variation in the average propensity to cite a given major constraint, in particular across countries, but also (to a lesser extent) across sectors.

Figure 2 below illustrates this variation for the "administrative regulation" constraint; in 2003, 15% of enterprises in France reported that administrative regulation was the major constraint on their business performance over the previous two years, whereas only 1% of enterprises in Portugal and Greece did so.

Figure 2: Proportions of SMEs in selected countries reporting administrative regulations as the major constraint on their business performance in 2003



Source: ENSR 2003 survey reported in KPMG et al. (2004b). Sample weighted to better reflect population distribution of enterprises. Sample of 7,459 SMEs; Liechtenstein and Iceland were in original sample but were omitted for brevity.

The empirical literature on small firm survival cited earlier suggests a range of structural and institutional characteristics that might produce sector and country-level effects on survival rates in Europe. These include returns to scale, advertising intensity, rapidity of innovation, availability of financing and sectoral growth rates. Macroeconomic effects on survival (such as changes in real exchange rates or unemployment) are unlikely to be identifiable in the short time period over which data are available. Firm-level factors such as educational qualifications of directors are likely to be averaged away in the sample, but

may have some influence on the fixed effects to the extent that there are major differences in the average incidence of such characteristics across the countries and sectors in the sample. There are also likely to be country-level policy effects that go beyond those captured by the survey evidence on perceived constraints.

Two effects suggested by theory can be incorporated directly using available data. These are the impact of entry on the survival prospects of earlier entrants and the effect of firm size on new entrant survival.

Several studies have reported a negative association between the rate of entry into a market in the current period and the survival prospects of firms that entered in earlier periods. This is related to another common finding that entry and exit rates tend to be highly correlated. Siegfried and Evans (1994), in a survey of past empirical work on entry and exit, note that the causality behind these results may work both ways. An increase in entry may depress survival rates as "existing firms are displaced by aggressive, more efficient new entrants."[143] However, there may also be a "vacuum effect" as business failures leave former customers without a source of supply, creating opportunities for new entry.[144] Indeed, there is a further possibility: the characteristics of firms that succeed in entering a market despite high entry barriers might be different from those that enter markets with low barriers. For example, surmounting high entry barriers might indicate that firms have higher productivity than other candidate entrants. If this is so, low firm birth rates might be an indicator of high average entrant quality, leading to good survival prospects.

Available data will not permit us to determine which of these effects is most important, but we include the small business birth rate in each period as an explanatory variable to capture the net effect on survival of displacement, vacuum and firm selection. If more data were available (particularly in the time dimension), it would be interesting to explore the vacuum effect further, e.g. by including lagged exit rates as an instrument for entry rates.

---

[143] Siegfried and Evans (1994), p.144.

[144] *Ibid*, p.147.

Another common empirical finding is that survival rates are positively associated with start-up size.[145] As per Audretsch (1995), to the extent that a sector is subject to increasing returns to scale, firms that enter at a larger scale should possess a cost advantage over smaller competitors. If firms learn about their potential profitability only after entry, this cost advantage could imply that larger entrants are on average less likely to fail. Size effects can be incorporated to a limited extent in the model by including dummy variables for each size band in the dataset (1-4, 5-9 and 10+ employees).

Regression models

The model posits that the probability of survival over a given term (e.g. one year) for a given firm in year *t*, country *c*, size band *s* and activity *a* is a function of its size (proxied by number of employees), the average birth rate ($\bar{b}$) of firms of its type during year *t*, the presence or absence of institutional constraints included in vector (**I**) and a range of sectoral, institutional, macroeconomic and firm level characteristics. These latter characteristics are proxied by sector, country and time dummies. The one-year probability of survival for firm *i* ($s1^i$) can be expressed as

$$s1^i = f\left(\bar{b}, \mathbf{I}, c, s, a, t\right) \tag{7}$$

Since only aggregate data are available, the model is transformed to one describing the one-year survival rate (*S*1) for all firms of a given type {*c*, *s*, *a*, *t*}:

$$S1_{csat} = \frac{\sum_i s1^i_{csat}}{N_{csat}} = g\left(\bar{b}, \bar{\mathbf{I}}, c, s, a, t\right) \tag{8}$$

where *N* is the total number of firms of a given type and the institutional constraints are measured by the average incidence of perceived constraints for the firm type. Two-year (*S*2) and three-year (*S*3) survival rates for each firm type are constructed in the same way.

[145] However, it is not an unchallenged result: Audretsch *et al.* (1999), using a probit model on Italian longitudinal data, found no evidence of a size effect on survival rates. Tveterås and Eide (2000) found evidence that size effects are less important for new small firms (such as those being modelled here) than for new plants of existing firms. McCloughan and Stone (1998) found that firm size at birth was not an important factor in survival of foreign-owned plants in Northern Ireland, although average size over a firm's lifetime was positively associated with survival.

Using sets of dummy variables for *c*, *s*, *a* and *t* and allowing for measurement error, a regression equation can be constructed for survival rates over any term. Given available data, 1, 2 and 3 year survival rates are of interest. These equations (omitting multiplicative coefficients) are shown below:

$$S1_{csat} = \alpha + BIRTHRT_{csat} + \mathbf{CONSTRAINTS}_{csat} + \mathbf{COUNTRY}_{c} + \mathbf{SIZE}_{s} + \mathbf{ACTIVITY}_{a} + \mathbf{TIME}_{t} + \varepsilon_{csat} \quad (9)$$

$$S2_{csat} = \alpha + BIRTHRT_{csat} + \mathbf{CONSTRAINTS}_{csat} + \mathbf{COUNTRY}_{c} + \mathbf{SIZE}_{s} + \mathbf{ACTIVITY}_{a} + \mathbf{TIME}_{t} + \varepsilon_{csat} \quad (10)$$

and

$$S3_{csat} = \alpha + BIRTHRT_{csat} + \mathbf{CONSTRAINTS}_{csat} + \mathbf{COUNTRY}_{c} + \mathbf{SIZE}_{s} + \mathbf{ACTIVITY}_{a} + \mathbf{TIME}_{t} + \varepsilon_{csat} \quad (11)$$

where $BIRTHRT_{csat}$ is equal to the cell mean new firm entry rate (firm births/total firms) for each firm type; and we include a classical disturbance term $\varepsilon_{csat} \sim N(0, \sigma^2) \ \forall \ c = 1,\ldots,C;\ s = 1,\ldots,S;\ a = 1,\ldots,A;$ and $t = 1,\ldots,T$. **CONSTRAINTS** is a vector containing the shares of firms within each type reporting a given constraint as the major one affecting business performance and **COUNTRY**, **SIZE**, and **ACTIVITY** are vectors of 1/0 dummy variables containing all but one category of each type in the data. When these models are estimated, the omitted categories in all cases are Constraint = "No major constraint", Country = Spain, Size = 1-4 staff and Year = 2002.

Expectations about coefficient values

Our main hypothesis is that coefficients on each of the constraint variables should be negative, since the omitted constraint category contains firms that report no major constraint. In particular, the subset of constraints that seem likely to materially increase firms' costs or limit their capacity for growth, including administrative regulations, access to finance and availability of skilled labour, should have a significant negative effect on survival rates.

We also expect that the coefficients on *BIRTHRT* will be negative. There may be added competitive pressure on recent entrants in markets where there are many subsequent entrants, and firms that successfully enter markets with high entry barriers (which might

well be correlated with the constraints we are measuring) might be "fitter" than those that enter markets with low entry barriers.

Recall that survival rates are likely to increase with size of firm. The two coefficients relating to size bands (empl_5_9 and empl_10p) should be positive, since the baseline contains the smallest firms. Moreover, the coefficient relating to the largest firms (with 10 or more employees) should be larger than the one for firms with 5-9 employees.

There is no theoretical basis for expecting a particular pattern of country or activity coefficients, except to say that there are likely to be significant differences among them reflecting omitted factors that have an impact on survival rates. Similarly, the time dummies will only be significant if there is a trend in survival rates of European micro firms generally.

Finally, we expect to see a decline in the average survival rate as firms' ages rise (i.e. from the one-year to the two- and three-year models). This should be reflected in the intercept terms, which should be positive, but less than 1.

Estimation strategy

The dependent variable in each of the regressions is fractional (i.e. survival rates fall in the interval [0,1]). OLS suffers from well-known shortcomings when applied to such data. Since some observations take a value of 1, the option of applying a logistic transformation to these variables is not available. However, the GLM quasi-likelihood estimator introduced in Papke and Wooldridge (1996) (hereafter referred to as "P-W") for use with fractional response data does accommodate 0 and 1 cases, and we include this as a second estimator to ensure that the results are robust.[146]

[146] Estimation was carried out in Stata 8, using the reg and glm commands, the latter with the following switches: family(binomial) link(logit) scale(x2).

7.3 Results

This section provides estimation results for the models set out in Section 7.2 above. The OLS and P-W results proved to be similar in both levels and significance of variables.

Separate regressions were carried out on the pooled data explaining one-, two- and three-year survival rates (i.e. Equations 9-11 above). Samples A, B and C respectively were used for these regressions, and the results are shown in Table 28 (using OLS) and Table 29 (using P-W) below. In the latter case, we show estimated marginal effects evaluated at the means of the independent variables in order to facilitate comparison with the OLS results. Regression results for this model are shown in Table 31, Annex 7.1.

Table 28: OLS regression results for new entrant survival rate

Dep Var.	***SURV1***		***SURV2***		***SURV3***	
	One-year survival rate		*Two-year survival rate*		*Three-year survival rate*	
Variables	**Coef.**	***t-stat.***	**Coef.**	***t-stat.***	**Coef.**	***t-stat.***
Constant	0.924	*84.46***	0.868	*43.14***	0.851	*25.94***
labour	-0.00294	*-0.27*	-0.0207	*-0.91*	-0.105	*-2.87***
finance	0.0149	*1.03*	-0.00907	*-0.35*	0.00338	*0.08*
newtech	0.00796	*0.37*	-0.0160	*-0.42*	-0.0413	*-0.55*
organisation	0.0361	*1.17*	-0.103	*-1.8**	-0.123	*-1.48*
qualmgmt	0.0310	*1.28*	-0.0129	*-0.32*	-0.0237	*-0.31*
regulation	-0.00773	*-0.59*	0.0178	*0.79*	-0.0395	*-0.94*
infrastruct	0.0207	*0.97*	-0.0934	*-1.88**	-0.133	*-1.83**
other	-0.0175	*-1.52*	-0.00430	*-0.21*	-0.133	*-3.4****
birthrt	-0.327	*-3.73****	-0.564	*-3.44****	-0.867	*-2.97****
dk	0.0271	*3.42****	-0.0230	*-1.52*	-0.0706	*-2.87****
fi	0.0426	*5.06****	0.0382	*2.48***	0.0416	*1.73**
it	0.0483	*6.74****	0.0269	*1.98***	0.0319	*1.61*
lu	0.0343	*4.99****	0.0240	*1.67**	-0.00795	*-0.42*
nl	-0.0573	*-7.53****	-0.117	*-6.47****		
no	-0.00172	*-0.21*	-0.0316	*-2.07***	-0.0285	*-1.09*
pt	0.0664	*8.67****	0.0831	*4.81****		
se	0.0646	*8.87****	0.0542	*3.27****	0.0580	*2.69****
uk	0.0240	*2.54***	-0.0654	*-3.84****		
y2000	0.00854	*1.72**	0.0177	*1.72**		
y2001	0.00844	*1.73**	0.0145	*1.41*	0.0282	*2.27***
construction	-0.00407	*-0.59*	-0.0275	*-2.12***	-0.0563	*-2.55***
wholesale	-0.00459	*-0.68*	-0.0209	*-1.76**	-0.0302	*-1.38*
retail	-0.00603	*-0.9*	-0.00654	*-0.56*	-0.0502	*-2.36***
hotels_cater	-0.0105	*-1.49*	-0.0226	*-1.61*	-0.0346	*-1.55*
transp_comms	-0.00989	*-1.49*	-0.0135	*-1.16*	-0.0436	*-2.02***
finserv	0.00287	*0.37*	0.00890	*0.63*	0.0214	*0.86*
businserv	-0.00198	*-0.3*	-0.00275	*-0.23*	-0.0132	*-0.62*
empl_5_9	0.0134	*2.66****	0.0430	*4.71****	0.0574	*3.67****
empl_10p	0.00666	*1.1*	0.0393	*3.45****	0.0653	*3.26****
Observations	493		364		234	
Adjusted R^2	0.498		0.497		0.339	
F-test (29,463)	17.8 [0.000]					
F-test (29,334)			13.4 [0.000]			
F-test (25,208)					5.78 [0.000]	
Heterosced-asticity	$\chi^2(1)$ = 2.87 [0.0900]		$\chi^2(1)$ = 0.25 [0.617]		$\chi^2(1)$ = 0.02 [0.889]	
RESET	F(3,460) = 5.29[0.0014]		F(3,331) = 1.76 [0.155]		F(3,205) = 0.94 [0.424]	
linktest hatsq	t = -1.22 [0.222]		t = -0.47 [0.637]		t = -0.02 [0.988]	

*Note: *, ** and *** denote significant at the 10%, 5% and 1% level respectively. Numbers in brackets are p-values. Heteroscedasticity test is Breusch-Pagan/Cook-Weisberg and specification tests are Ramsay RESET test and STATA's linktest. Data sources: see Table 27 above.*

Table 29: Papke-Wooldridge Estimator marginal effects results for new entrant survival rate, evaluated at sample means

Dep Var.	***SURV1***		***SURV2***		***SURV3***	
	One-year survival rate		*Two-year survival rate*		*Three-year survival rate*	
Variables	**Marg. Eff.**	***Z-stat.***	**Marg. Eff.**	***Z-stat.***	**Marg. Eff.**	***Z-stat.***
Constant	0.957	*10.52***	0.881	*9.96***	0.802	*8.04***
labour	-0.00451	*-0.46*	-0.0197	*-0.88*	-0.113	*-2.92***
finance	0.0145	*1.06*	-0.00487	*-0.18*	0.00329	*0.07*
newtech	0.00577	*0.29*	-0.0131	*-0.36*	-0.0532	*-0.68*
organisation	0.0418	*1.2*	-0.111	*-1.94**	-0.129	*-1.47*
qualmgmt	0.0277	*1.14*	-0.00454	*-0.12*	-0.0233	*-0.32*
regulation	-0.00514	*-0.45*	0.0200	*0.88*	-0.0421	*-0.96*
infrastruct	0.0139	*0.67*	-0.0912	*-1.95**	-0.141	*-1.89**
other	-0.0150	*-1.27*	0.00512	*0.23*	-0.144	*-3.42****
birthrt	-0.195	*-2.74****	-0.445	*-2.85****	-0.706	*-2.5***
dk	0.0153	*2.55***	-0.0223	*-1.44*	-0.0749	*-2.76****
fi	0.0269	*4.48****	0.0391	*2.79****	0.0469	*2***
it	0.0313	*6****	0.0251	*1.97***	0.0331	*1.69**
lu	0.0201	*4.11****	0.0214	*1.6*	-0.00796	*-0.42*
nl	-0.0328	*-4.81****	-0.107	*-5.14****		
no	-0.00164	*-0.25*	-0.0278	*-1.79**	-0.0293	*-1.07*
pt	0.0377	*7.01****	0.0785	*5.03****		
se	0.0436	*7.9****	0.0513	*3.41****	0.0560	*2.7****
uk	0.0121	*1.72**	-0.0593	*-3.27****		
y2000	0.00872	*1.77**	0.0160	*1.51*		
y2001	0.00816	*1.76**	0.0140	*1.33*	0.0300	*2.29***
construction	-0.00521	*-0.77*	-0.0290	*-2.08***	-0.0642	*-2.58****
wholesale	-0.00526	*-0.77*	-0.0221	*-1.74**	-0.0360	*-1.47*
retail	-0.00591	*-0.87*	-0.00785	*-0.64*	-0.0571	*-2.37***
hotels_cater	-0.0105	*-1.47*	-0.0246	*-1.6*	-0.0384	*-1.54*
transp_comms	-0.0103	*-1.51*	-0.0154	*-1.24*	-0.0513	*-2.1***
finserv	0.00197	*0.26*	0.00861	*0.59*	0.0209	*0.79*
businserv	-0.00384	*-0.58*	-0.00446	*-0.36*	-0.0168	*-0.72*
empl_5_9	0.0121	*2.66****	0.0418	*4.69****	0.0601	*3.88****
empl_10p	0.00709	*1.36*	0.0366	*3.48****	0.0644	*3.41****
Observations	493		364		234	
Log likelihood	-73.6		-97.2		-80.9	
AIC	0.420		0.699		0.914	
Specification: linktest hatsq	Z = 0.72 [0.474]		Z = 1.44 [0.150]		Z = 1.23 [0.218]	

*Note: *, ** and *** denote significant at the 10%, 5% and 1% level respectively. Standard errors scaled using square root of Pearson χ^2-based dispersion. Numbers in brackets are p-values. Specification test is STATA's linktest. Data sources: see Table 27 above.*

Baseline case and survival rate trend

As noted earlier, the omitted categories (and thus baseline case) for all three regressions describes the set of enterprises in 2002 that were based in Spain, had 1-4 staff, reported manufacturing as principal activity, and said that they faced no major constraint on business performance. The likelihood of a firm from this baseline group surviving the relevant number of years is represented by the intercept term in each model. These coefficients are highly significant and reflect the usual declining cumulative probability of post-entry survival: 95.7% for at least one year, 88.1% for at least two years and 80.2% for at least three years (using the P-W estimates). Of course, given that the dependent variable has an upper bound of 1, these intercept values imply that there is little variation in survival rates for our other variables to explain, particularly in the one-year survival model.

Constraint coefficients

As expected (since the baseline case reported no major constraint on performance), the constraint coefficients these coefficients tend to be negative. Indeed, all the constraint coefficients that approach statistical significance in any of the models are negative. A second encouraging result is that most of the constraint coefficients in the two-year model are smaller in absolute terms than their three-year counterparts. If perceived constraints represent real burdens on firms, it seems likely that they should have a cumulative effect over time.

However, few of the constraint effects are statistically significant. None is significant in the one-year model, and only "Lack of skilled labour" and "Other" are highly significant (both in the three-year model). The coefficients on "Infrastructure (road, gas, electricity, communication, etc.)" and "Implementing new forms of organisation" show similar patterns of marginal significance and rising absolute values in the two- and three-year models.

Surprisingly, classes of constraint that are given a great deal of attention in both policy circles and the academic literature do not approach significance in any of the models: "Administrative regulations (on environment, health and safety etc.)" and "Access to finance."

The magnitude of these effects can be illustrated by simulating the impact of a hypothetical change in one of the perceived constraints on firms in a particular country. For example, the estimates imply that if the share of firms in Luxembourg reporting "Lack of skilled labour" as their major constraint fell from its sample average level of 29.3% (highest in the sample) to the 18.5% level prevailing in Italy (lowest), and all of the affected firms moved to the "no major constraint" category, the three-year survival rate in Luxembourg should rise by about 1.4 percentage points, from 76.6% to 78.0%.[147]

Inspection of correlation matrices for the variables includes in these regressions shows that we need not be concerned about multicollinearity among the constraints variables; their pairwise correlations are low. To demonstrate this, the matrix for the variables in two-year survival rate regressions is shown in Table 32 in Annex 7.1 below.

Other results

The birth rate coefficients are highly significant and have the expected (negative) sign in all three models. The effect also seems to strengthen the longer a firm is in the market. This supports the prior expectation that current high rates of entry would be associated with lower post-entry survival prospects for earlier entrants.

Many of the country dummies are also significant, implying that national effects beyond those affecting entry rates and perceived constraints are important to survival rates. Sweden, Finland and Portugal are noteworthy in having significant positive coefficients in all models, while the Netherlands coefficient is negative in both years for which we have data. In most cases, the signs associated with individual country effects are broadly consistent across the three models; however, the Denmark, Luxembourg and UK dummies change sign between models.

Sector dummies seem to have less explanatory power than national ones. The construction sector shows significantly lower two- and three-year survival prospects than the baseline (manufacturing), and there is some evidence of specific negative effects for wholesaling, retailing and the transport and communications sector.

[147] This is based on the P-W results. For this illustration, we also assume that firm birth rates would not be affected by the improved availability of skilled labour. If they were (positively) affected, this would reduce the predicted improvement in survival rates.

As expected, there is evidence that survival rates increase with firm size. Enterprises with more than four employees seem to have about a 6-7% greater chance of surviving three years than smaller firms. This effect also seems to strengthen as duration from birth increases. There is no significant difference in the coefficients for firms with 5-9 employees and those with 10 or more, suggesting either that the size bands are too narrow to detect a difference or the size penalty is restricted to the smallest firms only.[148]

All the year dummies have positive coefficients, implying that 2000 and 2001 had lower survival rates than 2002, but levels of significance vary. It would probably require more time observations to detect substantial changes in survival rates from time-related factors such as changes in macroeconomic conditions.

The two- and three-year survival models pass specification tests even when estimated via OLS, but the one-year OLS model appears to be less robust.

[148] Equality of the empl_5_9 and empl_10p coefficients was not rejected. Using the P-W estimates, the results were $\chi^2(1) = 0.27$ [P=0.601] for the two-year survival model and $\chi^2(1) = 0.20$ [P=0.656] for the three-year model.

7.4 Conclusions and discussion

The average incidence of different perceived major constraints on small businesses' performance, as recorded by surveys, varies depending upon the country and sector surveyed. In this chapter, we have tested the hypothesis that types of small firms (defined by size, country and sector) reporting a higher incidence of major constraints should have worse survival prospects.

Our econometric models provide some evidence in support of this hypothesis, but it is not conclusive. Groups of firms reporting constraints have lower survival rates after controlling for other factors, and this effect seems to become more pronounced (both in magnitude and statistical significance) as the time since entry rises. However, the estimated effects are modest in absolute terms and appear to be restricted to only a few constraints included in the ENSR surveys; in particular, "Lack of skilled labour" and "Other". Surprisingly, we find no significant effects of administrative regulation and access to finance constraints on survival rates, despite strong priors on these variables.

While these results suggest that survey evidence on business perceptions captures some useful information on particular threats to post-entry performance, most perceived major constraints show no statistically significant relationship to survival rates, even over three years. Until more convincing evidence is available linking this type of survey evidence to outcomes, we should be cautious about using it in the assessment or design of public policy.

There seem to be possible explanations for why most perceived constraints seem to lack predictive power as to small firms' survival prospects.

First, firms' views on the main constraints they face may simply not be a good reflection of the institutional and other barriers to their survival. It would be surprising if firms under considerable stress could not perceive this, but it might be more plausible that such firms would find it hard to specify the precise sources and relative significance of their difficulties.

Second, perhaps the results are an accurate reflection of both the perceptions and the effects of major constraints. This would imply that the sorts of burdens on small businesses that tend to get the most attention in academic research and policy initiatives, excessive red tape and limited access to finance, may not be significant burdens at all. This could be either because they never were significant (which seems unlikely, given the

weight of existing theoretical and empirical evidence), or because the jurisdictions included in this sample have succeeded in mitigating their potential effects through policy. For example, excessive regulatory burdens on small firms may have been successfully offset through the use of thresholds or other policy measures. This explanation also seems unlikely, given the wide country-level variations in perceived regulatory constraints across the sample.

Finally, there may be a problem with the data. This seems the most likely explanation for our results. Two possible sources of error are worth considering.

One possibility is that the survey questions that were used may have failed to reveal the actual relationship between some perceived constraints and market conditions. Perhaps respondents are expressing a generalised level of competitive stress, rather than focusing on identifying the most onerous burden faced. This could explain why the regressions show the "right" signs but limited statistical significance for individual coefficients. It is also possible that the broad nature of the specified constraints or the backward-looking focus of the questions may have led to imprecise responses. A more serious potential problem (in the sense that it would be difficult to correct) is that of selection bias. For example, firms under serious pressure from excessive administrative regulation might be less likely to respond to surveys, which could tend to bias the regression coefficients on this constraint downwards.

Alternatively, the problem may lie with the set of firm types that remained after matching survey evidence with demographic data. We have already noted the limited variation in the dependent variable due to the availability of survival data up to only three years. Data availability may also have limited the variation in our perceived constraints variables. Although there was substantial variation across countries and sectors in these data, some of the most interesting countries could not be included in the analysis due to a lack of matching demographic data. For example, the countries shown in Figure 1 above with both the highest (France, Germany, Belgium) and lowest (Greece) tendency to report administrative regulation as the major constraint were omitted for this reason.

We conclude with some possible avenues for further research to help clarify if some aspect of the survey questions can account for these results or if one of the more fundamental explanations is correct.

It might be helpful to ask more specific questions about perceived constraints (e.g. about different sorts of regulations rather than administrative regulation in general). This could serve to reduce the "noise" associated with different firms interpreting the named constraints in different ways.

Another option is to ask survey questions about firms' current constraints, rather than backward-looking questions. One problem with relating backward-looking survey evidence to survival data is that it may impart a survivorship bias. For example, if all firms facing a particular major constraint were to exit the market, later backward-looking surveys would find few firms reporting this as a constraint. The apparent relationship of that constraint to survival rates would be biased downwards.

Asking questions that allow firms to indicate the perceived importance or intensity of particular burdens might be helpful.[149] The single major constraint formulation rules out interactions between different types of constraints and implicitly assumes that quite different types of burdens can be compared meaningfully by respondents. On the other hand, a question about the single major constraint is probably easier to answer than a series of more complex questions. An intermediate option might be to ask firms about a small number (e.g. 3) of most important constraints they face.

Most obviously, a broader set of outcomes data with a structure matched to that of the perceptions data would be useful (e.g. a longer time series, more disaggregated industrial classifications, and additional countries).

[149] For example, questions of this kind are asked in the IMD/World Economic Forum World Competitiveness Report and World Bank Investment Climate Surveys.

Annex 7.1: Further information on dataset

Table 30: Summary statistics for business constraints dataset

	Sample A (used in surv1 regression)				**Sample B** (used in surv2 regression)		**Sample C** (used in surv3 regression)	
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
surv1 surv2 surv3	0.946	0.0525	0.667	1	0.870	0.0813	0.795	0.0994
labour	0.202	0.201	0	1	0.178	0.182	0.238	0.210
finance	0.116	0.135	0	1	0.126	0.138	0.123	0.145
newtech	0.0380	0.0851	0	1	0.0412	0.0853	0.0364	0.0791
organisation	0.0210	0.0576	0	0.5	0.0223	0.0570	0.0243	0.0710
qualmgmt	0.0194	0.0734	0	1	0.0196	0.0807	0.0179	0.0767
regulation	0.0999	0.155	0	1	0.108	0.166	0.0940	0.164
infrastruct	0.0383	0.0846	0	1	0.0357	0.0651	0.0371	0.0808
other	0.200	0.187	0	1	0.213	0.195	0.156	0.187
noburden	0.266	0.222	0	1	0.257	0.224	0.274	0.218
birthrt	0.0396	0.0334	0	0.250	0.0371	0.0332	0.0367	0.0338
dk	0.0913	0.288	0	1	0.124	0.330	0.0983	0.298
es	0.132	0.339	0	1	0.0879	0.284	0.184	0.388
fi	0.0892	0.285	0	1	0.118	0.323	0.132	0.340
it	0.138	0.345	0	1	0.184	0.388	0.192	0.395
lu	0.130	0.336	0	1	0.121	0.326	0.179	0.385
nl	0.0913	0.288	0	1	0.0604	0.239	n/a	n/a
no	0.0690	0.254	0	1	0.0989	0.299	0.0726	0.260
pt	0.0872	0.282	0	1	0.0604	0.239	n/a	n/a
se	0.124	0.330	0	1	0.0824	0.275	0.141	0.349
uk	0.0487	0.215	0	1	0.0632	0.244	n/a	n/a
y2000	0.420	0.494	0	1	0.467	0.500	n/a	n/a
y2001	0.323	0.468	0	1	0.349	0.477	0.577	0.495
y2002	0.258	0.438	0	1	0.184	0.388	0.423	0.495
empl_1_4	0.371	0.484	0	1	0.387	0.489	0.397	0.490
empl_5_9	0.339	0.474	0	1	0.349	0.477	0.372	0.484
empl_10p	0.290	0.454	0	1	0.264	0.441	0.231	0.422
manufact	0.128	0.334	0	1	0.135	0.342	0.128	0.335
construction	0.132	0.339	0	1	0.110	0.313	0.137	0.344
wholesale	0.128	0.334	0	1	0.135	0.342	0.124	0.330
retail	0.128	0.334	0	1	0.140	0.348	0.128	0.335
hotels_cater	0.114	0.318	0	1	0.0852	0.280	0.120	0.325
transp_comms	0.134	0.341	0	1	0.143	0.350	0.124	0.330
finserv	0.0933	0.291	0	1	0.0962	0.295	0.0940	0.292
businserv	0.144	0.351	0	1	0.157	0.364	0.145	0.353

Source: analysis of data identified in Table 27.

Table 31: Papke-Wooldridge Estimator regression results for new entrant survival rate

Dep Var.	***SURV1***		***SURV2***		***SURV3***	
	One-year survival rate		*Two-year survival rate*		*Three-year survival rate*	
Variables	**Coef.**	***Z-stat.***	**Coef.**	***Z-stat.***	**Coef.**	***Z-stat.***
Constant	2.46	*10.52***	1.86	*9.96***	1.72	*8.04***
labour	-0.110	*-0.46*	-0.187	*-0.88*	-0.709	*-2.92***
finance	0.355	*1.06*	-0.0464	*-0.18*	0.0207	*0.07*
newtech	0.141	*0.29*	-0.125	*-0.36*	-0.3342	*-0.68*
organisation	1.02	*1.2*	-1.05	*-1.94**	-0.814	*-1.47*
qualmgmt	0.676	*1.14*	-0.0432	*-0.12*	-0.147	*-0.32*
regulation	-0.126	*-0.45*	0.191	*0.88*	-0.265	*-0.96*
infrastruct	0.339	*0.67*	-0.868	*-1.95**	-0.890	*-1.89**
other	-0.367	*-1.27*	0.0487	*0.23*	-0.903	*-3.42****
birthrt	-4.77	*-2.74****	-4.24	*-2.85****	-4.44	*-2.5***
dk	0.438	*2.55***	-0.200	*-1.44*	-0.427	*-2.76****
fi	0.902	*4.48****	0.420	*2.79****	0.317	*2***
it	1.05	*6****	0.254	*1.97***	0.217	*1.69**
lu	0.594	*4.11****	0.217	*1.6*	-0.0496	*-0.42*
nl	-0.629	*-4.81****	-0.788	*-5.14****		
no	-0.0394	*-0.25*	-0.245	*-1.79**	-0.176	*-1.07*
pt	1.54	*7.01****	1.06	*5.03****		
se	1.79	*7.9****	0.589	*3.41****	0.384	*2.7****
uk	0.338	*1.72**	-0.482	*-3.27****		
y2000	0.216	*1.77**	0.153	*1.51*		
y2001	0.206	*1.76**	0.135	*1.33*	0.187	*2.29***
construction	-0.122	*-0.77*	-0.256	*-2.08***	-0.373	*-2.58****
wholesale	-0.123	*-0.77*	-0.199	*-1.74**	-0.216	*-1.47*
retail	-0.1383	*-0.87*	-0.0732	*-0.64*	-0.334	*-2.37***
hotels_cater	-0.236	*-1.47*	-0.219	*-1.6*	-0.229	*-1.54*
transp_comms	-0.232	*-1.51*	-0.141	*-1.24*	-0.302	*-2.1***
finserv	0.0491	*0.26*	0.0841	*0.59*	0.136	*0.79*
businserv	-0.0910	*-0.58*	-0.0419	*-0.36*	-0.103	*-0.72*
empl_5_9	0.308	*2.66****	0.417	*4.69****	0.390	*3.88****
empl_10p	0.179	*1.36*	0.372	*3.48****	0.435	*3.41****

*Note: *, ** and *** denote significant at the 10%, 5% and 1% level respectively. Standard errors scaled using square root of Pearson χ^2-based dispersion. Numbers in brackets are p-values. Data sources: see Table 27 above.*

Table 32: Correlation matrix of variables used in two-year survival regressions (Sample B)

	surv2	birthrt	noburden	labour	finance	newtech	organisa-tion
surv2	1						
birthrt	-0.385	1					
noburden	-0.115	0.127	1				
labour	0.0413	-0.230	-0.223	1			
finance	-0.0283	0.0059	-0.260	-0.157	1		
newtech	0.0106	0.0154	-0.143	-0.107	-0.0542	1	
organisa-tion	-0.0038	-0.1086	-0.0911	-0.0283	-0.0546	0.0139	1
qual-mgmt	-0.0527	0.0372	-0.0867	-0.106	-0.0809	-0.0238	-0.0218
regula-tion	-0.0446	0.276	-0.224	-0.188	-0.144	-0.0976	-0.07
infra-struct	-0.0804	-0.0316	-0.123	-0.0123	-0.0615	-0.0361	-0.0894
other	0.1971	-0.154	-0.400	-0.286	-0.0491	-0.0355	-0.0297
es	-0.0483	0.033	0.0747	0.0079	0.0258	0.063	0.0705
dk	-0.0713	-0.0714	-0.0081	-0.0393	-0.055	-0.0368	-0.102
fi	0.142	-0.211	0.105	0.0711	-0.120	-0.0378	0.134
it	0.154	-0.202	-0.188	-0.057	0.247	-0.0009	0.0289
lu	0.0264	0.303	0.0584	0.0423	-0.132	0.0265	-0.0436
nl	-0.391	0.0609	0.0612	0.109	-0.0132	-0.0691	-0.01
no	-0.148	0.0511	0.114	-0.059	-0.0775	-0.0423	-0.0456
pt	0.259	0.0235	-0.208	0.0862	-0.0181	0.0534	-0.0495
se	0.204	-0.0506	0.047	-0.0414	0.0348	-0.0804	-0.0175
uk	-0.238	0.164	-0.0318	-0.0884	0.0622	0.147	0.0261
y2000	0.0907	0.0748	-0.092	-0.280	-0.0183	0.107	-0.0111
y2001	-0.121	-0.0302	-0.0273	0.191	0.0405	-0.104	-0.0488
y2002	0.0306	-0.0593	0.152	0.127	-0.0259	-0.0116	0.0738
manufac-turing	0.0816	-0.121	-0.054	0.0228	0.0184	-0.0063	0.0393
construc-tion	-0.0575	-0.0782	-0.0271	0.258	-0.0843	-0.0836	-0.0112
wholesale	-0.0779	-0.0671	-0.0037	-0.143	0.121	0.0076	0.0628
retail	0.0496	-0.124	0.04	-0.0514	-0.0128	0.0512	-0.0528
hotels_-cater	-0.028	-0.0007	-0.0728	-0.0077	0.112	-0.106	-0.117
transp_-comms	-0.0089	0.0009	0.0562	0.0199	0.0618	-0.0674	-0.0912
finserv	0.0224	0.264	0.125	-0.178	-0.198	0.126	0.0229
businserv	0.0076	0.144	-0.0619	0.0723	-0.028	0.0641	0.124
empl_1_4	-0.332	0.562	0.168	-0.256	0.0838	-0.0048	-0.086
empl_5_9	0.258	-0.250	-0.033	0.121	-0.0377	-0.0088	-0.0277
empl_10p	0.0901	-0.357	-0.153	0.155	-0.0528	0.015	0.127

	qual-mgmt	regula-tion	infra-struct	other	es	dk	fi
qual-mgmt	1						
regula-tion	-0.0797	1					
infra-struct	0.0054	-0.0635	1				
other	-0.0781	-0.206	-0.043	1			
es	0.0508	-0.1542	0.0905	-0.0785	1		
dk	-0.0511	0.179	0.0468	-0.0175	-0.120	1	
fi	-0.0647	-0.0273	-0.0892	-0.0425	-0.112	-0.135	1
it	0.0185	-0.106	0.0406	0.153	-0.151	-0.183	-0.171
lu	-0.0359	0.163	0.0579	-0.154	-0.117	-0.141	-0.132
nl	-0.0374	0.0299	0.0205	-0.144	-0.0807	-0.0976	-0.0913
no	0.176	-0.0363	-0.0039	-0.0308	-0.0988	-0.120	-0.112
pt	0.0081	-0.0852	0.0158	0.227	-0.0807	-0.0976	-0.0913
se	-0.0491	-0.0495	-0.104	0.0967	-0.0953	-0.115	-0.108
uk	-0.0098	0.0775	-0.109	-0.0241	-0.0826	-0.1	-0.0935
y2000	0.0646	0.0022	0.0013	0.301	0.0159	0.0108	0.107
y2001	-0.0101	0.0133	-0.0062	-0.118	-0.226	0.138	-0.0112
y2002	-0.0707	-0.019	0.0059	-0.244	0.254	-0.182	0.151
manufac-turing	-0.0152	-0.0192	0.0118	0.038	0.0455	-0.0048	0.0128
construc-tion	-0.0239	0.0003	-0.0956	-0.0645	-0.109	0.0328	0.0498
wholesale	-0.0237	-0.128	0.132	0.104	0.0546	0.0047	-0.031
retail	-0.045	-0.0671	0.0761	0.0539	0.0426	-0.0078	0.0098
hotels_ca ter	0.191	0.0839	-0.0947	-0.0273	-0.0953	-0.0241	-0.0119
transp_-comms	0.0289	-0.0479	0.0359	-0.0534	0.0369	-0.0137	0.004
finserv	-0.0464	0.183	-0.101	-0.0059	-0.0378	0.0447	-0.0279
businserv	-0.0364	0.0327	-0.0046	-0.051	0.0261	-0.025	-0.0071
empl_1_4	0.0505	-0.0253	0.0219	0.0032	-0.0135	-0.0488	0.0756
empl_5_9	-0.0466	0.0351	-0.0289	-0.036	-0.0236	-0.0297	0.113
empl_10p	-0.0057	-0.01	0.0069	0.0358	0.041	0.0871	-0.209

	it	**lu**	**nl**	**no**	**pt**	**se**	**uk**
it	1						
lu	-0.179	1					
nl	-0.124	-0.0951	1				
no	-0.151	-0.117	-0.0807	1			
pt	-0.124	-0.0951	-0.0659	-0.0807	1		
se	-0.146	-0.112	-0.0779	-0.0953	-0.0779	1	
uk	-0.127	-0.0974	-0.0674	-0.0826	-0.0674	-0.0797	1
y2000	-0.141	-0.0934	-0.244	0.114	0.27	0.0154	0.277
y2001	-0.0117	0.007	0.358	0.0026	-0.184	0.103	-0.189
y2002	0.196	0.112	-0.122	-0.150	-0.122	-0.145	-0.125
manufac-turing	-0.0046	0.002	-0.001	-0.0115	-0.001	-0.0331	-0.0055
construc-tion	0.043	-0.128	0.0246	0.0186	0.0246	0.0261	0.0202
wholesale	-0.0141	0.0114	0.0055	-0.0039	0.0055	-0.0264	0.001
retail	-0.0085	-0.001	-0.003	-0.014	-0.003	-0.0062	-0.0076
hotels_ca ter	0.0868	-0.112	0.0061	0.0461	0.0061	0.0536	0.0437
transp_-comms	-0.016	0.0664	-0.0071	-0.0187	-0.0071	-0.0396	-0.0116
finserv	-0.0383	0.109	-0.0064	-0.0378	-0.0064	0.0017	-0.01
businserv	-0.0304	0.0293	-0.0148	0.0261	-0.0148	0.0356	-0.0194
empl_1_4	-0.0373	-0.0005	-0.0409	0.0668	-0.0409	0.0645	-0.0259
empl_5_9	-0.0202	0.0185	-0.0162	-0.0855	-0.0162	0.0541	-0.0003
empl_10p	0.0638	-0.0196	0.0636	0.0185	0.0636	-0.131	0.0294
	y2000	**y2001**	**y2002**	**manufac-turing**	**construc-tion**	**wholesale**	**retail**
y2000	1						
y2001	-0.682	1					
y2002	-0.454	-0.342	1				
manufac-turing	0.0284	-0.0113	-0.0227	1			
construc-tion	0.0175	0.0208	-0.0479	-0.138	1		
wholesale	0.0195	-0.0289	0.0102	-0.154	-0.133	1	
retail	0.0043	-0.017	0.0152	-0.162	-0.140	-0.156	1
hotels_-cater	-0.0049	0.0385	-0.0406	-0.121	-0.105	-0.117	-0.123
transp_-comms	-0.0109	-0.0113	0.0278	-0.165	-0.143	-0.159	-0.167
finserv	-0.0305	0.0021	0.0367	-0.132	-0.114	-0.127	-0.134
businserv	-0.0245	0.0157	0.0123	-0.173	-0.149	-0.166	-0.175
empl_1_4	-0.0223	0.0252	-0.0021	-0.0235	-0.0009	0.0305	0.0033
empl_5_9	-0.0102	-0.0143	0.0305	0.016	0.0146	-0.0004	0.0269
empl_10p	0.0361	-0.0126	-0.031	0.0089	-0.0149	-0.0338	-0.033

	hotels_-cater	**transp_-comms**	**finserv**	**businserv**	**empl_1_4**	**empl_5_9**	**empl_10p**
hotels_-cater	1						
transp_-comms	-0.126	1					
finserv	-0.100	-0.137	1				
businserv	-0.131	-0.179	-0.143	1			
empl_1_4	-0.0389	-0.0097	0.0798	-0.0344	1		
empl_5_9	-0.0095	-0.0019	-0.0632	0.008	-0.592	1	
empl_10p	0.054	0.0129	-0.0205	0.0298	-0.475	-0.428	1

Note: Data described in Table 27 and summary statistics in Table 30 (Sample B) above.

References

Agarwal, R., 1998, "Small Firm Survival and Technological Activity", *Small Business Economics* 11, pp. 215–224.

Ahn, H. and Lee, M.-K., 1999, "An econometric analysis of the demand for access to mobile telephone networks", *Information Economics and Policy* 11, pp. 297–305.

Aidis, R., 2005, "Institutional Barriers to Small- and Medium-Sized Enterprise Operations in Transition Countries", *Small Business Economics* 25, pp. 305–318.

Aoki, R. and Small, J., 1999, *The Economics of Number Portability: Switching Costs and Two-Part Tariffs*, Working Paper, University of Auckland, November 1999.

Arellano, M. and Bond, S., 1991, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *The Review of Economic Studies*, 58(2), April, pp. 277-297.

Åstebro, T. and Bernhardt, I., 2003, "Start-up financing, owner characteristics, and survival", *Journal of Economics and Business* 55, pp. 303–319.

Atkinson, A. and Stiglitz, J., 1980, *Lectures on Public Economics* (McGraw-Hill, London).

Audretsch, D.B., 1991, "New Firm Survival and the Technological Regime", *The Review of Economics and Statistics* 73(3), August, pp. 441-450.

Audretsch, D.B., 1995, "Innovation, growth and survival", *International Journal of Industrial Organization* 13, pp. 441-457.

Audretsch, D.B. and Mahmood, T., 1994, "The rate of hazard confronting new firms and plants in U.S. manufacturing", *Review of Industrial Organization* 9(1), pp. 41-56.

Audretsch, D.B. and Mahmood, T., 1995, "New Firm Survival: New Results Using a Hazard Function", *The Review of Economics and Statistics* 77(1), February, pp. 97-103.

Audretsch, D.B., Santarelli, E., Vivarelli, M., 1999, "Start-up size and industrial dynamics: some evidence from Italian manufacturing", *International Journal of Industrial Organization* 17, pp. 965–983

Averch, H. and and Johnson, L., 1962, "Behaviour of the firm under regulatory constraint," *American Economic Review* 52(5), pp. 1053-1069.

Bain, J.S., 1951, "Relation of profit rate to industry concentration, American manufacturing: 1936-1940", *Quarterly Journal of Economics* 65, pp. 293-324.

Banerjee, A. and Ros, A., 2004, "Drivers of Demand Growth for Mobile Telecommunications Services: Evidence from International Panel Data" in Bohlin, E., Levin, S., Sung, N. and Yoon, C.-H. (eds.), *Global Economy and Digital Society* (Elsevier, Amsterdam).

Bartelsman, E., Scarpetta, S., and Schivardi, F., 2003, "Comparative Analysis of Firm Demographics and Survival: Micro-Level Evidence for the OECD Countries", *OECD Economics Department Working Papers*, No. 348 (OECD Publishing).

Bartlett, W. and Bukvič, V., 2001, "Barriers to SME Growth in Slovenia", *MOCT-MOST* 11, pp. 177-195.

Bernstein, M.H., 1955, *Regulating Business by Independent Commission* (Princeton University Press, Princeton, NJ).

Boeri, T. and Bellman, L., 1995, "Post-entry behaviour and the cycle: Evidence from Germany", *International Journal of Industrial Organization* 13, pp. 483-500.

Boyle, R., 2005, "Regulatory Impact Analysis: Lessons from the Pilot Exercise", CPMR Discussion Paper 31, Institute for Public Administration.

Brandt, N., 2004a, "Business Dynamics in Europe", *OECD Science, Technology and Industry Working Papers* 2004/1 (OECD Publishing).

Brandt, N., 2004b, "Business Dynamics, Regulation and Performance", *OECD Science, Technology and Industry Working Papers* 2004/3 (OECD Publishing).

Brunetti, A., Kisunko G., and Weder, B., 1998, *How Businesses See Government: Responses from Private Sector Surveys in 69 Countries*, Discussion Paper, International Finance Corporation, No. 33.

Buehler, S. and Haucap, J., 2004, "Mobile Number Portability", *Journal of Industry, Competition and Trade*, 4(3), pp. 223-238.

Buehler, S., Dewenter, R. and Haucap, J., 2006, "Mobile Number Portability in Europe", *Telecommunications Policy* 30(7), pp.385-399.

CMRG: Competition and Mergers Review Group, 1999, *Discussion in Relation to Groceries Order*, published by Ireland's Department of Enterprise, Trade and Employment, URL: http://www.entemp.ie/publications/commerce/2002/cmrg.pdf

Commission for Communications Regulation, 2004a, "Market Analysis - Wholesale Mobile Access and Call Origination, Consultation Paper", Document 04/05, URL: http://www.comreg.ie/_fileupload/publications/ComReg0405.pdf

Commission for Communications Regulation, 2004b, "Market Analysis – Wholesale Mobile Access and Call Origination, Response to Consultation and Notification to European Commission", Documents 04/118 and 04/118a. URL: http://www.comreg.ie/_fileupload/publications/ComReg04118.pdf

Demsetz, H., 1969, "Information and efficiency: another viewpoint," *The Journal of Law and Economics* 12, April, pp. 1-22.

Demsetz, H., 1973, "Industry Structure, Market Rivalry and Public Policy", *Journal of Law and Economics*, 16, pp. 1-10.

DeMuth, C.C., 1980a, "The White House Review Programs", *Regulation* 4(1), Jan/Feb., pp. 13-26.

DeMuth, C.C., 1980b, "The Regulatory Budget", *Regulation* 4(2), March/April, pp. 29-44.

Department of the Taoiseach, 2005a, *Report on the Introduction of Regulatory Impact Analysis*, July.

Department of the Taoiseach, 2005b, *RIA Guidelines: How to conduct a Regulatory Impact Analysis*, October. URL: http://www.betterregulation.ie/eng/index.asp?docID=78

ECC/CEPT: Electronic Communications Committee, 2005, "Implementation of Mobile Number Portability in CEPT Countries", ECC Report 31, European Conference of Postal and Telecommunications Administrations, Updated: October 2005, URL: http://www.ero.dk/documentation/docs/doc98/official/pdf/ECCREP031rev1.PDF

EIM/ENSR, 1995, *The European Observatory for SMEs: third annual report*, Zoetermeer: EIM Business and Policy Research.

Electronic Communications Appeal Panel, 2005, Decision No: 08/05 in respect of Appeal Numbers ECAP6/2005/03, 04, 05, 06, 07, 08. URL: www.ecap.ie.

Ellig, J., 2005, *Costs and Consequences of Federal Telecommunications and Broadband Regulations*, Working Paper, Mercatus Center, George Mason University, February.

EU Presidencies: Italian, Irish and Dutch Presidencies of the Council of the European Union, 2004, *A Comparative Analysis of Regulatory Impact Assessment in Ten EU Countries*, May, URL: http://www.betterregulation.ie/index.asp?docID=66.

European Council, 2004, Council Regulation No. 139/2004 on the control of concentrations between undertakings (the EC Merger Regulations), 20 January.

Eurostat, Business demography statistics, 2000-2002. URL: http://epp.eurostat.cec.eu.int/extraction/evalight/EVAlight.jsp?A=1&language=en&root=/theme4/sbs/bus_demo/size_cl

FCC: Federal Communications Commission, USA, 2004, *FCC Reports on Status of Local Number Portability*, Press Release, 13 May.

Fingleton, J., Evans, J. and Hogan, O., 1998, "The Dublin Taxi Market: Re-regulate or Stay Queuing?", *Studies in Public Policy* 3 (The Policy Institute: Dublin)

Freeman, R.B., 2002, *Institutional Differences and Economic Performance Among OECD Countries* (Centre for Economic Performance, London School of Economics and Political Science, London).

Gans, J. and King, S.P., 2000, "Mobile network competition, customer ignorance and fixed-to-mobile call prices", *Information Economics and Policy* 12, pp.301–327.

Gans, J.S., King, S.P. and Woodbridge, G., 2001, "Numbers to the people: regulation, ownership and local number portability", *Information Economics and Policy* 13, pp.167–180.

Greene, W., 2000, *Econometric Analysis*. Upper Saddle River, NJ: Prentice-Hall.

Hahn, R.W., 1989, "Economic Prescriptions for Economic Problems: How the Patient Followed the Doctor's Orders", *The Journal of Economic Perspectives* 3(2), Spring, pp. 95-114.

Hahn, R.W. and Hird, J., 1991, "The benefits and costs of regulation: review and synthesis." *Yale Journal on Regulation* 8, Winter, pp. 233-278.

Harhoff, D., Stahl, K., and Woywode, M., 1998, "Legal Form, Growth and Exit of West German Firms—Empirical Results for Manufacturing, Construction, Trade and Service Industries", *The Journal of Industrial Economics* 46(4), pp. 453-488.

Hausman, J.A., 1995, State Regulation of Cellular Prices", *Wireless Communications Forum* 3(1), April, pp. 61-68.

Hausman, J., 2002, "Mobile Telephone", Chapter 13 in Cave, M., Majumdar, S.K. and Vogelsang, I., *Handbook of Telecommunications Economics*: Volume 1 (North-Holland, Amsterdam).

Hausman, J. and Kuersteiner, G., 2004, *Difference in Difference Meets Generalized Least Squares: Higher Order Properties of Hypotheses Tests*, MIT economics Working Paper, June.

Hazlett, T. W. and Muñoz, R.E., 2004, "A Welfare Analysis of Spectrum Allocation Policies", AEI-Brookings Joint Centre, Related Publication 04-18, Aug 2004.

Hemmings, P. *et al.*, 2002, " The Role of Policy and Institutions for Productivity and Firm Dynamics: Evidence from Micro and Industry Data", *OECD Economics Department Working Papers*, No. 329 (OECD Publishing).

Honjo, Y., 2000, "Business failure of new firms: an empirical analysis using a multiplicative hazards model", *International Journal of Industrial Organization* 18, pp. 557–574.

Hopkins, T.D., 1997, "Alternative approaches to regulatory analysis: designs from seven OECD countries" in OECD, Regulatory Impact Analysis: Best Practices in OECD Countries.

http://www.cpb.nl/eng/pub/cpbreeksen/memorandum/136/memo136.pdf

IMF: International Monetary Fund, 2004, World Economic Outlook Database, September.

Infocomm Development Authority, Singapore, 2003, *Mobile Number Portability in Singapore*, Information Paper, August.

ITU: International Telecommunications Union, 2004, *World Telecommunication Indicators.*

Joskow, P.L. and Noll, R.G., 1981, "Regulation in theory and practice: an overview", in Fromm, Gary (ed.*), Studies in Public Regulation* (The MIT Press, Cambridge, Massachusetts).

Kaminski, J., "Dubliners still queueing for a living," *Irish Times*, 13 December 1999.

Kim, J., 2005, "The effect of mobile number portability on switching costs in the telecommunication industry", University of Wisconsin – Madison, Department of Economics Workshop, 19 October, URL: http://www.ssc.wisc.edu/econ/workshop/switching%20costs.pdf

Kingston University, Small Business Research Centre, 2005, *Regulation and Small Firm Performance and Growth: A Review of the Literature*, UK Small Business Service. URL: http://www.sbs.gov.uk/SBS_Gov_files/researchandstats/litreviewsmes.pdf

Klemperer, P., (1995) "Competition when consumers have switching costs: an overview with applications to Industrial Organisation, Macroeconomics, and International Trade", *The Review of Economic Studies* 62(4), pp. 515-539.

Kox, H., 2005, "Intra-EU differences in regulation-caused administrative burden for companies", Memorandum, CPB Netherlands Bureau for Economic Policy Analysis, December.

KPMG Special Services, EIM Business & Policy Research, and ENSR, 2002, *Business Demography in Europe, Observatory of European SMEs 2002*, No. 5, Luxembourg: Office for Official Publications of the European Communities. URL: http://europa.eu.int/comm/enterprise/enterprise_policy/analysis/doc/smes_observatory_2002_report5_en.pdf

KPMG Special Services, EIM Business & Policy Research, and ENSR, 2004a, "Highlights from the 2003 Observatory", *Observatory of European SMEs 2003*, No. 8, Luxembourg: Office for Official Publications of the European Communities. URL: http://europa.eu.int/comm/enterprise/enterprise_policy/analysis/observatory_en.htm

KPMG Special Services, EIM Business & Policy Research, and ENSR, 2004b, "General characteristics (Part A)", *Observatory of European SMEs 2003*, Luxembourg: Office for Official Publications of the European Communities. URL: http://europa.eu.int/comm/enterprise/enterprise_policy/analysis/doc/ensr_2003_tables/tables_survey_2003_general_parta.pdf

Laffont, J.-J., and Tirole, J., 1993, *A theory of incentives in procurement and regulation* (The MIT Press, Cambridge, Massachusetts).

LECG, 2004, *Study of Competition in the Provision of Non-investment Banking Services in Ireland: Report and Recommendations*, 29 November, published by Ireland's

Competition Authority. URL: http://www.tca.ie/professions/banking/banking_study.pdf

Lee, S-W., Kim, D.J. and Park, M.-C., 2004, "Demand for Number Portability in the Korean Mobile Telecommunications Market: Contingent Valuation Approach", *Proceedings of the 37th Hawaii International Conference on System Sciences – 2004*.

Levine, M.E. and Forrence, J.L., 1990, "Regulatory capture, public interest and the public agenda: toward a synthesis", *Journal of Law, Economics and Organisation* 6, pp.167-198.

Levy, D., 1985, "Specifying the Dynamics of Industry Concentration," *Journal of Industrial Economics* 34(1), September, pp. 55-68.

Loayza, N.V. *et al*., 2004, Regulation and Macroeconomic Performance, *The World Bank Policy Research Working Paper Series*, No. 3469.

Madden, G. and Coble-Neal, G. (2004). Economic determinants of global mobile telephony growth. *Information Economics and Policy* 16, pp. 519–534.

Mahmood, T., 2000, "Survival of Newly Founded Businesses: A Log-Logistic Model Approach", *Small Business Economics* 14, pp. 223–237.

Majone, G., 1996, *Regulating Europe* (Routledge, London).

Martin, S., 1994, *Industrial Economics: Economic Analysis and Public Policy* (Prentice-Hall, New Jersey).

Lenard, T.M. and Mast, B.D., 2003, "Taxes And Regulation: The Effects Of Mandates On Wireless Phone Users", *Progress on Point* 10, The Progress and Freedom Foundation, 18 October.

Mata, J. and Portugal, P., 1994, "Life Duration of New Firms", *Journal of Industrial Economics* 27(3), pp. 227–246.

Mata, J. and Portugal, P., 2002, "The Survival of New Domestic and Foreign-Owned Firms", *Strategic Management Journal* 23, pp. 323–343.

Mata, J., P. Portugal and Guimarães, P., 1995, "The survival of new plants: Start-up conditions and post-entry evolution", *International Journal of Industrial Organization* 13, pp. 459-481.

McCloughan, P. and Stone, I., 1998, "Life Duration of Foreign Multinational Subsidiaries: Evidence from UK Northern Manufacturing Industry 1970-93", *International Journal of Industrial Organization* 16(6), November, pp. 719-47.

McCubbins, M.D., Noll, R.G. and Weingast, B.R., 1987, "Administrative procedures as instruments of political control," *Journal of Law, Economics and Organisation* 3(2), Fall, pp. 243-77.

McNamara, G. and O'Hara, J., 2005, "Internal Review and Self-Evaluation; the Chosen Route to School Improvement in Ireland?", *Studies In Educational Evaluation* 31, pp. 267-282.

Merrill Lynch, 2004, Global Wireless Matrix, Q2 2004.

Milgrom, P. and Roberts, J., 1992, *Economics, Organization and Management* (Prentice-Hall, London)

Monopolies and Mergers Commission (UK), 1995, *Telephone number portability: A report on a reference under section 13 of the Telecommunications Act 1984*, November.

National Audit Office, 2005, *Evaluation of Regulatory Impact Assessments Compendium Report 2004-05*, 17 March.

NERA: National Economic Research Associates, 2003, *Switching Costs*, Office of Fair Trading and Department of Trade and Industry Economic Discussion Paper 5, Annexe C, case studies, April.

NERA-Smith: National Economic Research Associates and Smith System Engineering, 1998, *Feasibility Study & Cost Benefit Analysis of Number Portability for Mobile Services in Hong Kong, Final Report for OFTA*, May.

OECD, 1997,*Regulatory impact analysis: best practices in OECD countries* (OECD Publishing).

OECD, 2001a, *Regulatory Reform in Ireland* (OECD Publishing).

OECD, 2001b, *Businesses' Views on Red Tape: Administrative and Regulatory Burdens on Small and Medium-sized Enterprises* (OECD Publishing).

OECD, 2005, Quarterly National Accounts Database.

OFT: Office of Fair Trading, 2002, *Guidelines for competition assessment: A guide for policy makers completing Regulatory Impact Assessments*, February.

OFT: Office of Fair Trading, 2003a, *The control of entry regulations and retail pharmacy services in the UK: a report of an OFT market investigation (revised edition)*, March.

OFT: Office of Fair Trading, 2003b, *The private dentistry market in the UK*, March.

OFT: Office of Fair Trading, 2003c, *The regulation of licensed taxi and PHV services in the UK*, November.

OFTA: Office of the Telecommunications Authority, Hong Kong, 2006, *Mobile Number Portability in Hong Kong*. URL: http://www.ofta.gov.hk/en/ca_bd/mnp.html

Oftel, 1997, *Economic Evaluation of Number Portability in the UK Mobile Telephony Market*, July.

Oftel, 2003, "Review of competition: mobile access and call origination", a consultation issued by the Director General of Telecommunications, 11 April.

Ogus, A., 1998, "Regulatory Appraisal: A Neglected Opportunity for Law and Economics," *European Journal of Law and Economics* 6, pp. 53-68.

OIRA: Office of Information and Regulatory Affairs, US Office of Management and Budget, 2005, *Validating Regulatory Analysis: 2005 Report to Congress on the Costs and Benefits of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities.* URL: http://www.whitehouse.gov/omb/inforeg/2005_cb/final_2005_cb_report.pdf

Ovum, 2000, *Mobile numbering and number portability in Ireland: A report to the ODTR*.

Ovum, 2005, Mobile *Number Portability - an international benchmark, A report to MTN*.

Padilla, A.J., 1995, "Revisiting Dynamic Duopoly with Consumer Switching Costs", *Journal of Economic Theory* 67, pp. 520-530.

Papke, L. E. and Wooldridge, J., 1996, "Econometric methods for fractional response variables with an application to 401(k) plan participation rates", *Journal of Applied Econometrics* 11, pp. 619–632.

Peltzman, S., 1981, "Current developments in the Economics of Regulation", in Fromm, Gary (ed.), *Studies in Public Regulation* (The MIT Press, Cambridge, Massachusetts).

Persson, H., 2004, "The Survival and Growth of New Establishments in Sweden, 1987–1995", *Small Business Economics* 23, pp. 423–440.

Pierre, G and Scarpetta, S., 2006, "Employment protection: Do firms' perceptions match with legislation?", *Economics Letters* 90, pp. 328–334.

Purcell, D., 2004, "Competition and regulation in the retail pharmacy market", *Studies in Public Policy* 14 (The Policy Institute, Dublin).

Radaelli, Claudio M., 2005, "Diffusion without convergence: how political context shapes the adoption of regulatory impact assessment," *Journal of European Public Policy* 12(5), October, pp. 924-943.

Radaelli, Claudio, M., 2004, "The diffusion of regulatory impact analysis – Best practice or lesson-drawing", *European Journal of Political Research* 43(5), pp. 723-747.

Regulatory Impact Unit, 2003, "Better Policy Making: A Guide to Regulatory Impact Assessment", UK Cabinet Office.

Rose, N.L., 1985, "The incidence of regulatory rents in the motor carrier industry," *Rand Journal of Economics*, 16(3), Autumn, pp. 299-318.

Samuelson, L. and Zhang, J., 1992, "Search costs and prices", *Economics Letters* 38, pp. 55-60.

Schmalensee, R., 1987, "Horizontal merger policy: problems and changes", *Economic Perspectives* 1(2), pp. 41-54.

Schmalensee, R., 1989, "Intra-industry Profitability Differences in US Manufacturing 1953-1983", *Journal of Industrial Economics* 37, June, pp. 337-357.

SCM Network, 2005, *The International SCM Manual; measuring and reducing administrative burdens for businesses*, October. URL: http://www.administrative-burdens.com/filesystem/2005/11/international_scm_manual

Shew, W., 1994, "Regulation, Competition and Prices in Cellular Telephony", Working Paper No. 4329, The American Enterprise Institute for Public Policy Research.

Shy, O., 1995, *Industrial Organization: Theory and Applications* (The MIT Press, London)

Siegfried, J.J. and Evans, L.B., 1994, "Empirical Studies of Entry and Exit: A Survey of the Evidence", *Review of Industrial Organization* 9, pp. 121-155.

Smyth, J., 2006, "European court upholds ban on foreign pharmacists", *The Irish Times*, 14 July.

Spiller, P.T., 1996, "Institutions and Commitment", *Industrial and Corporate Change* 5(2), pp. 421-454.

Spiller, P.T. and Tiller, E.H., 1997, "Decision costs and the strategic design of administrative process and judicial review", *Journal of Legal Studies* 26(6), pp. 347-370.

Spulber, D.F., 1989, *Regulation and Markets* (The MIT Press, Cambridge, Massachusetts).

Stigler, G., 1971, The Theory of Economic Regulation, *Bell Journal of Economics and Management Science* 2(1), Spring, pp. 3-21.

Storey, D.J., 1994, *Understanding the Small Business Sector* (Routledge, London and New York).

Storey, D.J. and Wynarczyk, P., 1996, "The survival and non-survival of micro firms in the UK", *Review of Industrial Organization* 11, pp. 211-229.

Telegeography, 2005a, “Operators’ opinions split over MNP”, *CommsUpdate*, 22 September. URL: http://www.telegeography.com/cu/article.php?article_id=9159&email=html

Telegeography, 2005b, “Date set for wireless number portability”, *CommsUpdate*, 22 December. URL: http://www.telegeography.com/cu/article.php?article_id=10423&email=html

TIO: Telecommunications Industry Ombudsman, Australia, 2002, *Annual Report*.

Treasury Board of Canada, 1997, “Regulatory Reform Through Regulatory Impact Analysis: The Canadian Experience”, *Managing Better* 14, Evaluation, Audit and Review Group, February.

Tveterås, R. and Eide, G. E., 2000, “Survival of New Plants in Different Industry Environments in Norwegian Manufacturing: A Semi-Proportional Cox Model Approach”, *Small Business Economics* 14, pp. pp. 65–82.

US Census Bureau, 2005, International Database, URL: http://www.census.gov/ipc/www/idbsprd.html

Valletti, T.M., and Cave, M., 1998, “Competition in UK Mobile Communications”, *Telecommunications Policy* 22(2), pp. 109-131.

Viard, V.B., "Do Switching Costs Make Markets More or Less Competitive?: The Case of 800-Number Portability", *Rand Journal of Economics*, forthcoming.

Wallsten, S.J., 2001, “An Econometric Analysis of Telecom Competition, Privatisation, and Regulation in Africa and Latin America”, *Journal of Industrial Economics* 49(1), pp.1-19.

Walsh, P.P., 2004, “Quantifying the Impacts of Regulation”, Better Regulation Conference, Royal Hospital Kilmainham, Dublin, 12 May.

Wooldridge, J. M., 2002,. *Econometric Analysis of Cross Section and Panel Data* (The MIT Press, Cambridge, Massachusetts)

World Bank, Doing Business database, Washington DC. URL: http://rru.worldbank.org/doingbusiness

World Bank, 2004, *World Development Indicators*.

Index

D

E

F

G

H

I

J

K

L

M

N

O

P

R

S

T

U

V

W