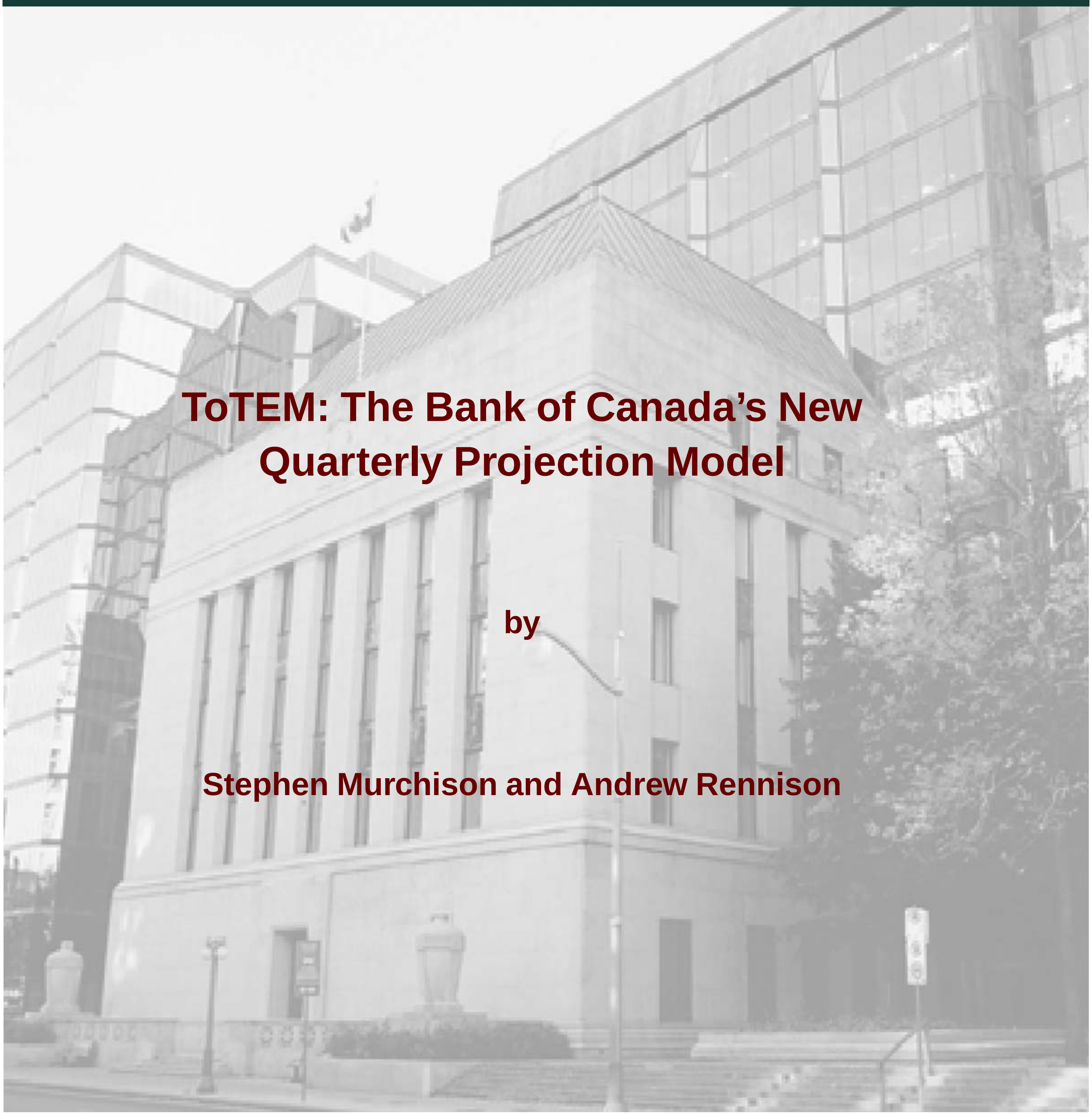
Bank of Canada Banque du Canada

Technical Report No. 97 / Rapport technique nº 97



# ToTEM: The Bank of Canada's New Quarterly Projection Model

by

Stephen Murchison and Andrew Rennison

December 2006

# ToTEM: The Bank of Canada's New Quarterly Projection Model

Stephen Murchison and Andrew Rennison

Research Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9

ISSN 0713-7931

Printed in Canada on recycled paper

# Contents

# Acknowledgements

The ToTEM project benefited from several important contributions. For their work in model building, we thank André Binette, Patrick Perrier, and Zhenua Zhu. We thank Amy Corbett, Debbie Gendron, Hope Pioro, and Manda Winlaw for excellent technical assistance. The authors benefited greatly from comments and discussion at the Central Bank Modelling workshops in Amsterdam in 2003 and Montréal in 2004, particularly from Klaus Schmitt-Hebbel and Alex Wolman. Valuable input was also received at Bank of Canada workshops and seminars, as well as in informal discussions with Robert Amano, Hafed Bouakez, Pierre Duguay, Rochelle Edge, Claude Lavoie, Andrew Levin, Rhys Mendes, Louis Phaneuf, Nooman Rebei, and Gregor Smith. Finally, we wish to express our gratitude to Don Coletti for his leadership and enthusiasm during the construction and implementation of the model.

# Abstract

The authors provide a detailed technical description of the Terms-of-Trade Economic Model (ToTEM), which replaced the Quarterly Projection Model (QPM) in December 2005 as the Bank's principal projection and policy-analysis model for the Canadian economy. ToTEM is an open-economy, dynamic stochastic general-equilibrium model that contains producers of four distinct finished products: consumption goods and services, investment goods, government goods, and export goods. ToTEM also contains a commodity-producing sector. The behaviour of almost all key variables in ToTEM is traceable to a set of fundamental assumptions about the underlying structure of the Canadian economy. This greatly improves the model's ability to tell coherent, internally consistent stories about the current evolution of the Canadian economy and how it is expected to evolve in the future. In addition, ToTEM's multiple-goods approach enables the Bank to gain insight into a much wider variety of shocks, including relative-price shocks. In particular, ToTEM is better equipped to handle terms-of-trade shocks, such as those stemming from movements in world commodity prices. But ToTEM does not mark a radical departure from QPM's design philosophy; rather, it should be regarded as the next step in the evolution of open-economy macro modelling at the Bank. Indeed, ToTEM adopts most of the features that distinguished QPM from its predecessors, including a well-defined steady state, an explicit separation of intrinsic and expectational dynamics, an endogenous monetary policy rule, and an emphasis on the economy's supply side. However, ToTEM extends this basic framework, allowing for optimizing behaviour on the part of households and firms, both in and out of steady state, in a multi-product environment.



*JEL classification: E17, E20, E30, E40, E50, F41*
*Bank classification: Economic models; Business fluctuations and cycles*

# Résumé

Les auteurs exposent en détail les caractéristiques techniques du modèle TOTEM (pour ***T**erms-**o**f-**T**rade **E**conomic **M**odel*), qui a remplacé le Modèle trimestriel de prévision (MTP) en décembre 2005 à titre de principal modèle utilisé par la Banque du Canada pour l'analyse de politiques et l'établissement de projections relatives à l'économie canadienne. TOTEM est un modèle d'équilibre général dynamique et stochastique adapté à un cadre d'économie ouverte. Il comprend quatre catégories de produits finis : les biens et services finaux, les biens d'équipement, les biens publics et les biens d'exportation. Il compte aussi un volet distinct pour le secteur des produits de base. Le comportement de la quasi-totalité des variables clés de TOTEM a son origine dans un ensemble d'hypothèses fondamentales concernant la structure de l'économie canadienne. Le modèle se trouve ainsi beaucoup mieux à même de décrire de manière cohérente et logique l'évolution actuelle — et anticipée — de l'économie canadienne que ne l'était le MTP. Par ailleurs, la structure à produits multiples de TOTEM permet à la Banque d'étudier les effets d'un éventail de chocs bien plus large, y compris les chocs de prix relatifs. En particulier, TOTEM se prête mieux à l'examen des variations des termes de l'échange, notamment de celles qui découlent des fluctuations des cours mondiaux des produits de base. Toutefois, le modèle TOTEM ne rompt pas avec les principes qui ont présidé à la conception du MTP; il marque simplement un nouveau jalon dans l'évolution des travaux de modélisation en économie ouverte à la Banque. TOTEM reprend en effet la plupart des éléments qui distinguaient le MTP des modèles qui l'ont précédé, dont un régime permanent bien défini, la différenciation explicite entre la dynamique intrinsèque du modèle et la dynamique liée aux anticipations, une règle de politique monétaire endogène et l'accent mis sur le secteur de l'offre. Cependant, TOTEM va plus loin en permettant de prendre en compte le comportement optimisateur des ménages et des entreprises dans une économie à produits multiples, aussi bien en régime permanent que hors équilibre.

*Classification JEL : E17, E20, E30, E40, E50, F41*
*Classification de la Banque : Modèles économiques; Cycles et fluctuations économiques*

# Introduction

In December 2005, staff at the Bank of Canada began using a new model for projections and analysis of the Canadian macroeconomy. This new model, called ToTEM (**T**erms-**o**f-**T**rade **E**conomic **M**odel), replaces the **Q**uarterly **P**rojection **M**odel (QPM), which had served in this role since its introduction in the early 1990s. When QPM was first introduced, it was considered state of the art among central bank models. It possessed a steady state that was well grounded in economic theory, it accounted explicitly for both stocks and flows, and it assumed an important role for forward-looking behaviour on the part of firms, households, and the central bank. Since then, advances in economic modelling, combined with enormous increases in computing power, have led to a new generation of macroeconomic models that build on the basic design philosophy underlying QPM. ToTEM incorporates several of the innovations that have emerged from the macroeconomic modelling literature, and thus brings the Bank's workhorse model back to the leading edge of central bank modelling. In essence, ToTEM takes advantage of the technological progress in economic modelling and computing power that has occurred over the past decade to enhance the fundamental strengths of QPM. The new model has a stronger theoretical foundation, is easier to work with, and better explains the dynamics of the Canadian economy.

As its description suggests, the Bank's quarterly projection and policy-analysis model must fulfill two main objectives. First, it is used to produce the Bank staff's quarterly economic projection of the Canadian economy, which is an important input into the monetary policy decision-making process at the Bank of Canada. The model's role in the projection implies that large weight is placed on its ability to explain the observed features of Canadian economic data (Coletti and Murchison 2002 and Macklem 2002). Second, it is used as a tool to examine a wide range of policy-related questions, including issues related to optimal monetary policy and the interpretation and implications of a wide range of shocks. As stressed by Lucas (1976), any examination of the interaction between policy and economic outcomes requires a model whose parameters are invariant to the behaviour of policy. In short, the model is required to distinguish between dynamics generated by the structure of the economy and those generated via agents' expectations, which in turn requires strong theoretical underpinnings. Meeting the requirements of these dual objectives—that the model be both theory-based and data consistent—has, in the past, proven a difficult task. Indeed, when QPM was built, many features that were inconsistent with the theoretical core were added to the model, since the theory upon which QPM's core was based did not provide adequate explanations for the observed dynamics of the business cycle.

In the past 10 years, a great deal of progress has been made on this front, as a result of a push among academics and central bank modellers to develop and test better theory-based models of business-cycle dynamics. Advances in the theoretical literature have come as a result of a convergence between two once-competing lines of economic research. This so-called "new neoclassical synthesis" (NNS) combines the neoclassical real business-cycle framework of rational optimizing agents with the New Keynesian framework of market imperfections, such as nominal rigidities and imperfect competition, to create a unified theory of business-cycle analysis.[1] An explosion of subsequent research has used models based on the fundamentals of the NNS, often referred to as dynamic stochastic general-equilibrium (DSGE) models, to explain various aspects of business cycles.

At the same time, increased computing power has facilitated the application of more sophisticated econometric techniques to large-scale macro models. Recently, formal estimation and empirical evaluation of DSGE models have yielded encouraging insights into their potential ability to explain the broad features of macroeconomic data and to provide reliable forecasts. A variety of techniques, from the moment-matching techniques in Christiano, Eichenbaum, and Evans (2005, hereafter CEE), to the Bayesian maximum-likelihood estimation technique employed by Smets and Wouters (2005), have been used to evaluate these models' empirical coherence, and the results have been encouraging.

The structure and implementation of ToTEM have benefited greatly from these recent advances. In particular, the core theory underlying ToTEM—that of utility-maximizing households and profit-maximizing firms with rational expectations—has been augmented with several fairly recent modelling innovations that deserve special mention here.

[1] This term was coined by Goodfriend and King (1997), who provide a comprehensive discussion of this line of research.

For households, we assume a functional form for utility in which labour and leisure are not additively separable (King, Plosser, and Rebelo 1988), which ensures that, under any values for the preference parameters, labour supply will be stationary in the presence of trend productivity growth. Stationary labour supply is a feature that is apparent in the aggregate data but inconsistent with more conventional additively separable specifications of utility. Our utility function also includes internal habit formation, as in CEE, which implies that households care about their level of consumption relative to their lagged consumption. The assumption of habit persistence smooths the response of consumption to movements in, for example, real interest rates.

As is typical in the DSGE literature, manufacturing firms sell their goods in monopolistically competitive markets and thus set prices above marginal cost. Their production technology is a constant elasticity of substitution (CES) aggregate of commodities, effective labour and capital services, and imports. Sticky prices and nominal wages are modelled in a modified Calvo (1983) set-up in which some portion of firms (workers) is randomly selected each period to reoptimize their price (wage), while the remaining firms index their price (wage) to an index of lagged and steady-state price (wage) inflation. As in CEE, variable capital utilization is possible, but comes at a cost in terms of foregone production, and changes in capital and investment are subject to quadratic adjustment costs. Variable capital utilization smooths the response of marginal cost to movements in production, while adjustment costs on investment allow the model to produce a gradual response of investment to movements in the cost of capital.

In place of the conventional capital rental market assumption, in which capital can be costly reallocated across firms, we assume that capital is firm-specific, as in Altig, Christiano, Eichenbaum, and Linde (2004) (hereafter ACEL) and Woodford (2005). In the more typical model of perfect capital mobility, a firm's marginal cost is invariant to the level of demand for its good. By contrast, when capital is owned by the firm and quasi-fixed in the short run, firm-level marginal cost is increasing in its output. Overall, the assumption of firm-specific capital reduces the sensitivity of prices to marginal cost, thereby making the model's predictions consistent with the observed sensitivity of aggregate inflation to demand conditions while at the same time allowing for average price contract durations that accord with micro survey evidence.

Of course, a model for Canada would not be complete without a role for international trade and an adequate accounting of Canada's role as an exporter of commodities. In this regard, ToTEM combines the NNS with an open-economy, multi-goods framework that allows for international trade in finished products and commodities.[2] The model contains producers of four distinct finished products: consumption goods and services, investment goods, government goods, and export goods. ToTEM also contains a commodity-producing sector. Imports are treated as an input to production following McCallum and Nelson (2000). An imported good makes its way to market as follows: an importing firm buys imports from the foreign economy and subsequently sells them to a manufacturing firm at a price that is temporarily fixed in the domestic currency. We model sticky import prices in the same framework as that for domestic-goods prices and wages, and since both imported-input prices and final-goods prices are sticky, the model possesses an element of vertical or supply-chain price staggering, which is crucial to generating realistic exchange rate pass-through.

A commodity producer combines a fixed factor, which we call land, with capital services and labour, and sells this good to domestic and foreign distributors in a perfectly competitive market. Domestic commodity distributors then sell commodities to domestic manufacturers or consumers directly at a sticky price analogous to those for imports. The latter assumption ensures that commodity-price movements affect consumer prices more gradually than if we assumed manufacturers bought commodities directly from the commodity producer.

Our discussion of ToTEM will proceed as follows. In chapter 1, we provide the background for the development of ToTEM, including a brief discussion of the role that models have traditionally played at the Bank of Canada and that ToTEM will be required to fill. Some historical context is provided, including a discussion of ToTEM's predecessor, QPM, and the advances that established that model at the forefront of central bank modelling. We then discuss the areas in which we believe ToTEM represents an improvement over QPM. Chapter 2 provides the formal derivation of the model from the model's theoretical framework through the behavioural equations that arise from the optimization of rational firms and households. To aid in intuition, the model's key linearized equations are also presented. Chapters 3 and 4 discuss the model's parameterization and simulation properties, respectively. Chapter 5 offers some conclusions.

[2] As such, ToTEM belongs to the family of open-economy DSGE models, or "new open economy models" (NOEM). Lane (2001) provides an excellent survey of the new open-economy literature.

# Chapter 1

# ToTEM's Origin and Purpose

## 1.1 Models at the Bank of Canada

Models have long played a central role in policy discussions at the Bank of Canada.[1] Models are valued at the Bank because of the structure they impose on the economic debate. They provide a framework within which to view economic developments, and force analysts to formalize views that may be largely based on intuition. Disagreement about, for example, the economic forecast can be separated into the assumptions regarding those factors determined outside the model and differences in view about the structure of the economy, as summarized by the model. Models can also be used to settle disagreements that cannot be resolved with theory alone. Specifically, while theory often tells us what forces are at play in determining economic outcomes, these forces are often offsetting; a model helps determine the relative importance of each factor and can thus provide an estimate of their net impact.

Bank staff prefer to use several economic models to guide policy, rather than just one. This approach is taken for two reasons. First, there is uncertainty regarding the correct economic paradigm (Jenkins and Longworth 2002 and Selody 2001). In economics there is no laboratory in which economists can alter one key variable at a time and then directly observe its impact on the economy. As a result, there is considerable debate in academic and policy circles about which paradigm

[1]For a thorough discussion of the history of model building and use at the Bank of Canada, see Duguay and Longworth (1998). For a more recent discussion of the role of models at the Bank, see Coletti and Murchison (2002).

best represents the way in which, for example, monetary policy affects inflation. By using several models based on competing paradigms, Bank staff help guard against serious policy errors that could result from relying on a single paradigm that may be incorrect.

Second, models differ depending on their intended purpose, and no single economic model can answer all questions that are relevant for monetary policy. For example, a purely statistical model will do well as a short-term forecasting device, but will typically fail to identify the underlying equilibrating forces in the economy. The latter are particularly important for monetary policy, which requires a medium-term perspective and, thus, a clearer representation of the economy's equilibrium. Over a longer-term horizon, therefore, the usefulness of a purely statistical model for monetary policy tends to diminish. In addition, as we will discuss later, the expectations-formation process is particularly important for monetary policy, because it is one of the key channels through which monetary policy affects economic outcomes. Purely statistical forecasting models are silent on the role of expectations and are therefore of limited usefulness when considering questions related to monetary policy.

That said, the staff rely most heavily on one main model for constructing macroeconomic projections and conducting policy analysis for Canada. This workhorse model reflects the consensus view of the key macroeconomic linkages in the economy. The interrelationships among aggregate variables, most notably output, inflation, interest rates, and the exchange rate, are viewed as particularly important, and the structure and parameterization of the central model therefore pay them special attention. The focus on macro properties implies that more detailed questions often arise that the main model is not designed to answer. In these cases, insights from more specialized models are often overlaid on the main analysis. But it is through the workhorse model that economic events are interpreted by Bank staff.

For the past 12 years, the Bank's main Canadian model has been the Quarterly Projection Model. But the Bank's tradition of developing and using macroeconomic models dates back much further. Modelling at the Bank of Canada began in earnest in the mid-1960s with RDX1 (Research Department Experimental), a fairly simple Keynesian model of price and income determination. It was supplanted shortly after its completion by RDX2, a much larger and more detailed model that contained, among other things, an extensive financial sector. This model was used as both a learning tool and as a means of analyzing the medium-term economic outlook. But the detailed structure of RDX2 implied that simulations were often difficult to understand. As well, the monetarist revolution of the 1970s implied that RDX2, with its Keynesian focus on demand-side disturbances, was not always well suited for policy discussions of the day. Together, these factors prompted economists at the Bank of Canada to develop new, more transparent models for policy analysis. These models were used to consider specific questions related to monetary policy, such as the interaction between inflation and money growth, but were not detailed enough to be considered seriously as forecasting tools. A noteworthy example was SAM (Small Annual Model), which was developed in the early 1980s and was used to address issues of a more medium-term nature. While SAM was small by macroeconomic-model standards (about 25 equations), its structure was complex and its short-run dynamic properties were considered at odds with the generally accepted view of the transmission mechanism of monetary policy. Nevertheless, the experience gleaned from building a more theoretical model such as SAM would have an important influence on the development of QPM at the end of the 1980s.

In the late 1970s, Bank staff began conducting regular model-based forecast exercises and, to this end, RDXF was developed. RDXF represented a means for collecting and coordinating the input of Bank staff, with ownership of the model distributed among analysts familiar with a particular sector. As will be discussed below, this bottom-up approach led to a decentralized sense of model evaluation, in which the fit of individual equations was considered ahead of the model's predictions for the properties of the aggregate economy. Furthermore, in the mid- to late 1980s, there was an increasing appetite within management at the Bank for longer-term simulations that would provide insight into the unwinding of the rising inflation, rising interest rate environment that existed in Canada at that time. In the late 1980s, the Bank decided to undertake a project to build a new model that could be used both as a forecasting tool and to analyze the effects of monetary policy in a meaningful way.

## 1.2 The experience with QPM

The result of this project was QPM, which the Bank adopted as its Canadian projection model in September 1993. QPM's design philosophy was motivated largely by the need to address the deficiencies associated with previous-generation macroeconomic models. Its structure reflected a desire to abstract from sector-specific details of the economy in order to focus on the core macro linkages in a theoretically consistent framework. Instead of the bottom-up approach that characterized previous forecasting models at the Bank and elsewhere, its builders placed the primary focus on the model's overall simulation properties. Because of this emphasis on system properties rather than sectoral details, QPM was relatively small compared with other central bank models.

Before QPM, models used at the Bank often did not possess a long-run equilibrium to which simulations would eventually converge, and therefore they were unable to meaningfully address medium- to long-term issues. For example, the typical early model did not account for stock variables such as government debt. As a result, the model would predict that a deficit-financed tax cut would lead to stronger consumption in the short term, but would ignore the associated negative implications of an increase in debt. In some of the worst cases, model simulations often cycled or drifted without limit when faced by certain economic shocks. Given the forward-looking nature of the projection exercise at the Bank, and the need to quickly produce basic risk analyses, it became clear that such models were inadequate. Accordingly, QPM was designed to be dynamically stable around a well-defined steady state.

Another important shortcoming of 1970s and 1980s macro models was the primitive way in which they accounted for agents' expectations. In fact, if expectations played any role in these models, they were often modelled as purely adaptive. Moreover, the model's parameters were typically estimated under the assumption that they were independent of the policy regime and therefore stable across regimes. Lucas (1976) argues, however, that if, in reality, expectations were formed rationally, the parameters of these "reduced-form" equations would depend on agents' expectations of policy, and would therefore not likely be stable if there were important changes in policy regime.

Similarly, model simulations conducted under policy regimes that were not reflective of the chosen sample could produce highly misleading results. For Canada, the time-series properties of inflation are an important example. Phillips curves estimated for the 1970s tend to support the claim that inflation drifts without an anchor rather than returning to its average value. Since the early 1990s, however, the degree of inflation persistence has fallen dramatically (see Longworth 2002 for a survey of the empirical literature for Canada), suggesting that inflation tends to return fairly quickly to its long-run average value (the inflation target) following a shock. One interpretation of the change in the behaviour of inflation is as follows: first, expectations of future inflation are an important determinant of inflation today; second, these expectations take into account the Bank's efficacy in achieving its objectives; and third, the Bank has increased the transparency with which it conducts monetary policy since the 1970s. If, as this explanation suggests, the time-series properties of inflation depend crucially on the monetary authority, it is critically important to properly characterize policy and expectations when conducting model simulations.

QPM was built such that the link between monetary policy and expectations plays a key role in shaping economic outcomes. In doing so, the monetary authority conditions the expectations of agents such that they are consistent with this goal. In the current policy regime, the monetary authority reacts to shocks in such a way as to achieve the 2 per cent over the medium term. By committing to act in this manner, the monetary authority ensures that expectations of inflation remain anchored to the target. This is accomplished through the incorporation of an endogenous policy rule in the model. This approach is both conceptually and operationally distinct from the previous practice of imposing an exogenous path for the instrument of policy for each simulation.

These early models also emphasized a detail-oriented "bottom-up" approach, often consisting of hundreds of unrestricted equations. More importantly, however, they were typically built on an equation-by-equation basis, with little emphasis placed on the model-simulation properties. Sector specialists estimated so-called "demand" or "price" equations that could typically be interpreted only in a partial-equilibrium setting. As a result, problems relating to econometric identification emerged, and implausible exclusion restrictions were required to identify the parameters of the model. As a result, when model equations were put together in a system, they often produced unrealistic simulation properties. In addition, the high level of detail that resulted from following this approach, combined with a general lack of theory-based restrictions, led to a proliferation of unrestricted model parameters, which in turn led to large standard errors and poor performance in out-of-sample forecasting.

Finally, previous-generation models typically ignored the supply side of the economy. This omission was likely based on the then-prevailing Keynesian view that most disturbances to output were driven by fluctuations in demand. In reality, experience has taught us that supply disturbances also play an important role in explaining history. Moreover, supply disturbances have fundamentally different consequences for prices, wealth, and potential output than do demand shocks. Taking supply disturbances seriously is a feature of modern macroeconomics, and this was reflected in QPM.

For more than 12 years, QPM was the workhorse Canadian projection and policy-analysis model at the Bank. The model has shaped the way the staff think about the economy and the shocks that hit the Canadian economy. It has also helped us understand some of the key Canadian macroeconomic issues of the 1990s, such as the impact of government debt (Macklem, Rose, and Tetlow 1995), the effect of central bank credibility on optimal monetary policy (Amano, Coletti, and Macklem 1999), the advantages of price-level versus inflation targeting (Maclean and Pioro 2000), and the costs and benefits of price stability (Black, Coletti, and Monnier 1997). QPM has also had a major impact on the modelling efforts of other inflation-targeting central banks including the Reserve Bank of New Zealand and the Swedish Riksbank, both of whom employ variations of QPM. More recently, QPM has significantly influenced modelling efforts at the Bank of Japan.

## 1.3 The new model

In the 12-plus years since the first QPM-based projection, significant advances have been made in the field of macroeconomic modelling. Foremost among these advances are improved techniques for modelling the dynamics of the business cycle from a theoretical perspective, a better understanding of the determinants of inflation and the interaction between inflation and monetary policy, and an increase in computing power and improvements in solution techniques that have facilitated not only the use, but also the formal econometric evaluation, of much larger models.[2] For the most part, however, these advances build on the basic foundations that were established by models such as QPM in the early 1990s. In this sense, ToTEM should be regarded as the next step in the evolution of open-economy macro modelling at the Bank of Canada, rather than as a radical departure from QPM's design philosophy. Indeed, most of the features that distinguished QPM from its predecessors, including a well-defined steady state, the explicit separation between intrinsic and expectational dynamics, an endogenous monetary policy rule, and an emphasis on the economy's supply side, are all present in ToTEM. Unlike QPM, however, ToTEM incorporates optimizing behaviour on the part of firms and households, both in and out of steady state. It has become feasible to have the same optimizing decision rules that define the model's behaviour in the short and medium term as those that pin down its long-run steady state. Consequently,

[2]Since the outset of the project, AIM (see Anderson and Moore 1985, Anderson 1997), in conjunction with the LKROOTS command in TROLL, has been used to numerically linearize and solve ToTEM. The ability to do fast numerical linearizations greatly increases the variations of model features we are able to experiment with. In addition, AIM as implemented in TROLL has proved to be an extremely flexible, powerful tool.

the model's dynamics can always be traced back to a set of fundamental assumptions about the structure of the economy. This increased reliance on economic theory in the dynamic model, in turn, results in model simulations that are easier to understand and explain.

In what follows, we discuss the specific ways in which the practice of macroeconomic model building has evolved since QPM was built, and how these advances are reflected in ToTEM. At a conceptual level, three main features distinguish ToTEM from QPM. First, there exist multiple production sectors in ToTEM, each consisting of an assumed production technology, a set of demand functions for the various inputs to production, and a pricing schedule. In contrast, QPM implicitly assumed the existence of a single good, which was intended to represent GDP. The second difference is the explicit assumption that firms and consumers behave in an optimizing manner, both in the steady-state and dynamic versions of the model. In other words, firms seek to maximize their profits, while consumers/workers seek to maximize their happiness, which is captured by a "utility" function in consumption and leisure. The third difference is the assumption of rational, but not necessarily perfect foresight, expectations in ToTEM. We discuss each of these conceptual differences in detail.

## 1.4 A multi-goods framework

One area in which large advances have been made recently in the macro modelling literature is in the operationalization of larger-scale, multiple-sector models. In the past, the time required to solve forward-looking general-equilibrium models implied a significant trade-off between the degree of detail contained in a model and its ability to inform policy in real time. But given the combination of linearization/solution techniques and the cheap computational power available to modellers, the burden implied by the extra detail of a multiple-sector framework is no longer a major consideration.

A multi-goods framework is particularly useful for issues related to open-economy macroeconomics. Indeed, in a one-good optimization-based paradigm, there is no rationalization for the bilateral trade of goods between countries; a country either exports or imports its good. Also, in a one-good model, consumption baskets in the home and foreign country are by definition identical, and therefore the law of one price must hold. As such, there is no explanation for fluctuation in the real exchange rate. The specification of multiple sectors creates a motivation for trade and thus allows modellers to consider important questions such as those related to real exchange rate and current account determination, all in the context of an optimization-based framework.

Models with sectoral detail can also be used to consider questions such as the cause and effect of domestic relative-price movements. At first glance, it may seem that such detail is of second-order interest for a monetary-policy maker, but consider that the Bank of Canada currently targets the rate of inflation of the consumer price index (CPI). It is reasonable to assume that the impact that a movement in aggregate demand will have on CPI inflation will differ, depending on whether it is the result of an increase in, for example, consumption or investment demand. In a one-good framework, a 1 per cent increase in aggregate demand will have roughly the same initial impact on prices whether it is the result of an increase in investment or consumption demand. On the other hand, in a model with separate production sectors, the source of the shock is very important. A positive investment demand shock will have no immediate impact on marginal cost in the consumption-goods-producing sector; it will affect consumption prices only to the extent that it causes factor prices, most notably wages, to increase. One could certainly envision circumstances where investment and consumption move in opposite directions, and therefore this distinction would become extremely important.

Furthermore, in a multi-goods set-up, one can consider issues such as the best inflation rate to target from the perspective of welfare maximization. For example, it may be the case that the current practice of targeting consumer-price inflation, which is made up of both imported and domestically produced goods, is suboptimal from a welfare perspective. Obviously, questions such as this cannot be fully addressed within the one-good framework.

The theoretical core of QPM was limited to a one-good paradigm for a number of reasons. First, at the time QPM was built, multi-sector general-equilibrium modelling was in its infancy; multi-sector specifications were typically limited to small, highly stylized models designed to deal with very specific issues. Also, as mentioned, the models that preceded QPM at the Bank of Canada put a large emphasis on capturing the partial-equilibrium behaviour of the economy's various sectors, and it is likely that the decision to restrict the theoretical core of QPM to a single sector was partly a response to the inability of the previous-generation models to produce reasonable macro simulation properties. More generally, the reality at the time of QPM's construction was that the maintenance of a general-equilibrium multi-sector model carried a prohibitively high cost in terms of computational burden. It was decided instead to make simpler additions to the model's core that, although theoretically incompatible with the model's core theory, could still enhance the model's ability to deal with the issues for which the core theory was not well suited.

Several features that were inconsistent with the one-good paradigm were layered on the core of QPM. First, to differentiate exports and imports, fixed import shares were specified for each component of GDP. The level of exports was then calculated residually, given both the level of imports and the level of net exports that were necessary to support the net foreign asset position.

To consider relative-price movements, auxiliary Phillips curves were added to the model for the various prices. Not only were these changes at odds with the model's one-good framework, they did not adequately capture the impact that relative-price movements could have on real variables. For example, terms-of-trade shocks, and commodity-price shocks more specifically, had no direct effect on the level of GDP. Also, real commodity prices had no discernible impact on the real exchange rate, whereas reduced-form estimates suggest that these effects are quite large for persistent commodity-price movements (see Amano and van Norden 1995).

Since QPM was built, multi-sector models have become more common. Macklem's (1992,1993) terms-of-trade model is an early example; while founded on the optimization-based framework of earlier, more stylized, models, it contains enough detail to produce informative dynamic simulations for shocks, such as movements in the terms of trade. A more recent example is the IMF's Global Economic Model (GEM, see International Monetary Fund 2004), a multi-country, multi-sector optimization-based model that has been used to examine, among other things, the determinants of international exchange rate fluctuations.

Given these advances in modelling techniques and computing speed, it was decided that ToTEM would also be based upon this multi-goods framework. ToTEM separates production into the main components of the national accounts: consumption, investment, government, and commodity and non-commodity exports. Heterogeneous production and pricing in these sectors is characterized by their distinct import concentrations, levels of technology, and levels of demand. In the future, these distinctions may also include different capital-to-labour ratios in, for example, the production of government goods versus the production of consumer goods. This separation provides theoretical motivation for the existence of relative prices for the various components of the national accounts.

Relative to QPM, this multi-sector set-up expands the set of questions that can be handled within the model's theoretical framework. Discussions about shocks such as movements in relative prices can be considered without relying on atheoretical add-on equations or judgment gleaned from other models. At the same time, ToTEM retains the overall emphasis on the economy's broad macro linkages that was a key consideration when QPM was built.

### 1.4.1 *A separate role for commodities*

ToTEM distinguishes between commodity-producing firms and producers of finished products. This separation is important for Canada, not only because roughly 11 per cent of GDP in Canada represents production of raw commodities,[3] but also because the finished-products and raw materials sectors are characterized by different technologies and competitive structures. Commodity production is subject to a number of real rigidities and, as such, supply is highly price inelastic in the short run. At the same time, it is difficult to differentiate a commodity produced in Canada from one produced abroad. Thus, the industry is much closer to perfect competition, at least from the perspective of a Canadian producer, than is the production of finished products where product differentiation is common. To properly understand the effects of commodity-price shocks, it is therefore necessary that a model contain an explicit distinction between the commodity-producing and manufacturing sectors as well as their respective markets.[4]

Given the one-good set-up assumed in QPM, commodities played no explicit role in the model, and commodity prices were treated as an add-on in the price equations. Specifically, commodity prices were assumed to influence directly the price of exports (and the price of imports to a much lesser extent) and, as a result, Canada's terms of trade. The principal shortcoming of this set-up, at least in terms of its aggregate implications, was that real GDP and the real exchange rate were counterfactually insensitive to commodity-price movements.

In contrast, commodities play a critical role in the theoretical core of ToTEM. Indeed, by incorporating commodities in the model, several insights were gleaned

[3] Value added in agriculture, fishing and trapping, mining, and a number of resource-based manufacturing industries (wood products, paper and allied products, primary metal products, petroleum and coal products, and chemical products) as a percentage of GDP, in 1997 constant dollars.

[4] Bank staff are currently developing a version of ToTEM that will allow a separate role for energy and non-energy commodities.

from previous optimization-based models of the Canadian economy. Two important examples include the MACE model (Macro and Energy model, see Plourde 1987) and Macklem (1992, 1993). Like these models, ToTEM includes both a commodity-producing sector and a role for commodities in the production of the finished product. Finally, ToTEM goes one step further, allowing commodities to also serve as a finished good that is purchased directly by consumers, an assumption that in turn allows a theory-based distinction between core and total consumption prices.

## 1.5 Micro foundations with optimizing agents

While QPM went partway towards satisfying the recommendations of the Lucas critique, ToTEM carries the optimizing-agent/rational-expectations framework essentially throughout the model. Dynamic decision rules for all variables in ToTEM reflect utility- and profit-maximizing decisions by rational agents, whereas QPM's reliance on the optimization-based framework was limited to the core of the steady state and therefore did not influence the model's short- and medium-term dynamics. This was largely due to the fact that, on its own, the steady-state model's optimizing structure was unable to explain the business-cycle fluctuations in the macroeconomic data. For this reason, it was decided that additional structure would be built around the theoretical core to allow the model to explain the business-cycle regularities required for short-term forecasting.

Of the features added to dynamic QPM, the most prominent involved the addition of lags of variables to the model's equations that could not be justified on the basis of theory. In some cases, lags were justified on the basis that they represented adaptive expectations; in other cases, no interpretation was given. Other features include the addition of lags of the output gap to QPM's Phillips curve equations, a modification intended to impart a causal relationship between the output gap and inflation, which was consistent with the staff's view of the monetary transmission mechanism. As a result of this approach, none of the parameters that shaped the model's short- and medium-run dynamics was traceable to the original theoretical framework.

Since QPM was built, however, advances in the modelling of market imperfections—real and nominal rigidities—have resulted in fewer atheoretical modifications being required to replicate the business-cycle features of the data. These advances allow ToTEM to maintain a unified structure; that is, the optimizing decision rules that define the model's behaviour in the short and medium term are the same as those that pin down its long-run steady state. The increased reliance on economic theory in the dynamic model results in model simulations that are easier to understand and to explain. For example, the degree to which inflation responds to movements in marginal cost can be directly traced to assumptions about the degree of competition in the economy, while the sensitivity of consumption to real interest rates is traceable to consumers' underlying tastes. Below, we describe three concrete examples of differences that emerge between ToTEM and QPM because of the assumption of optimizing behaviour on the part of firms and consumers.

### 1.5.1 *Real marginal cost: A unified framework of price determination*

The pricing equations or Phillips curves in QPM included a role for the lagged output gap, which is defined as the difference between current output and the level "consistent with an unchanged rate of inflation over the short run" (Butler 1996). The principal theoretical motivation for including the output gap was as follows: markets may be regarded as being perfectly competitive in the long run, which implies that price is equated with marginal cost at the firm level. But over the business cycle, firms whose products are in high demand will tend to enjoy some monopoly pricing power, which will cause them to choose a price that is higher than marginal cost (a positive markup). Therefore, at an aggregate level, desired margins will be procyclical over the business cycle and zero in the long run. The builders of the QPM then opted to use the output gap as a proxy for these unobserved desired margins. In this sense, the output gap introduces a time-varying wedge between price and marginal cost. In addition, the view was taken that the firm-level short-run supply curve was essentially flat, so a firm's productivity and marginal cost was independent of its production level (treating the price of factor inputs such as the wage as given). However, this latter assumption, in particular, is inconsistent with the structure of QPM, whose production technology implies substantial variation in a firm's productivity over the business cycle. Moreover, theoretical assumptions that would give rise to a procyclical desired markup (for example, non-constant elasticity of demand) did not carry through to the rest of the model, creating numerous inconsistencies.

In ToTEM, by contrast, marginal cost drives inflation not by assumption, but rather because firms set prices to maximize profits in an environment where the elasticity of demand for goods, and thus firm's desired markup, is invariant to the state of the economy. As a result, inflation is driven by expected future movements in marginal cost. Marginal cost is increasing in firm-level output for several reasons, and therefore the short-run supply curve is upward sloping. For instance, higher production will tend to be associated with higher capital investment and more intensive utilization of existing capital. Both of these factors reduce productivity at the firm level, given that the installation of new capital causes production disruptions and higher rates of utilization cause the existing capital stock to depreciate faster.

Nevertheless, while the channels through which excess demand creates higher inflation are somewhat different (an upward-sloping short-run supply curve in ToTEM versus procyclical desired markups in QPM), both models predict qualitatively the same behaviour for inflation for all of the key shocks. Quantitative differences typically have more to do with how ToTEM was calibrated than with differences in assumed market structure.

### Price determination in ToTEM and QPM

To illustrate the main differences between ToTEM and QPM in terms of how prices are affected by economic developments, consider a simplified model of firm pricing behaviour whereby the short-run production technology for firm $i$ is given by

$$Y_{it} = L_{it}^{\alpha}, \tag{1.1}$$

so that the firm can increase its output only by increasing labour input, $L$, in period $t$. In a world where firms are free to reset their price in every period, the assumption of profit maximization implies that firms will set prices as a markup over marginal cost, and that markup will depend on the degree of competition in the firm's product market,

$$p_{it} = \mu_{it}\lambda_{it}, \tag{1.2}$$

where $p_{it}$ denotes the firm's relative price ($p_{it} = P_{it}/P_t$), $\lambda_{it}$ denotes the firm's real marginal cost of production, and $\mu_{it}$ is the markup. Given the production technology in equation (1.1), combined with the assumption that firms take the aggregate real wage as given, the marginal cost of production will be

$$\lambda_{it} = \frac{w_t L_{it}}{\alpha Y_{it}}. \tag{1.3}$$

If we again use equation (1.1) to eliminate labour from our definition of marginal cost and then log-linearize the resulting expression, we have

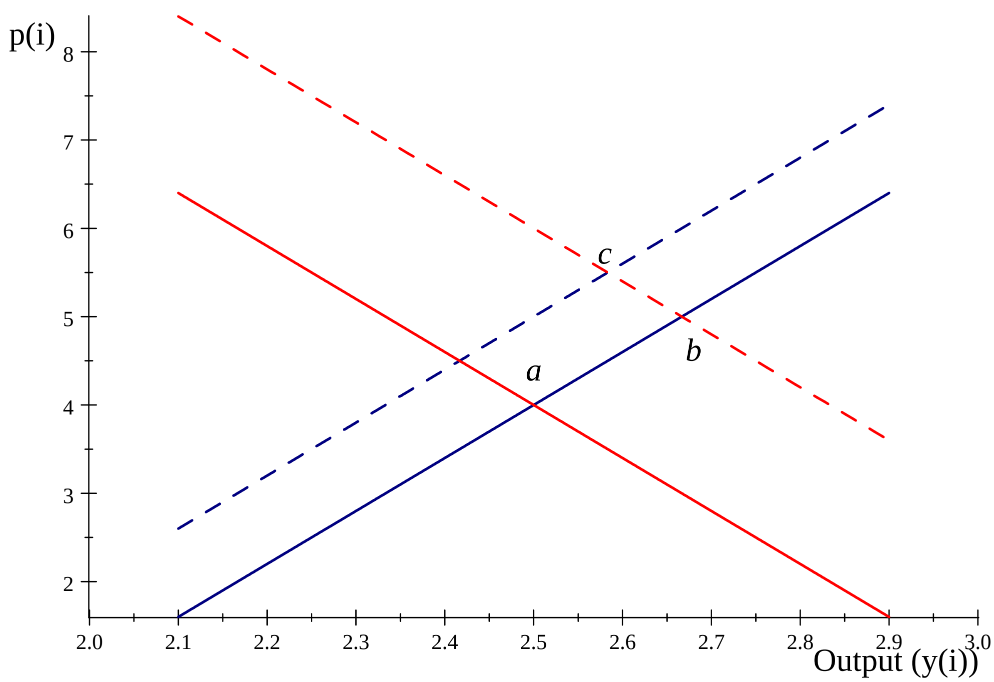
$$\widehat{\lambda}_{it} = \widehat{w}_t + \left(\frac{1-\alpha}{\alpha}\right)\widehat{Y}_{it}, \tag{1.4}$$

where the $\widehat{\ }$ denotes per cent deviation from some arbitrary point. Therefore, given our first-order condition (1.2), the firm's supply curve will be

$$\widehat{Y}_{it} = \left(\frac{\alpha}{1-\alpha}\right)\left(\widehat{p}_{it} - \widehat{w}_t - \widehat{\mu}_{it}\right), \tag{1.5}$$

which we display in two-dimensional $\left\{\widehat{p}_{it}, \widehat{Y}_{it}\right\}$ space below.

**Figure 1: Demand/Supply in ToTEM**




In our simple model, then, there are just two variables that can cause the short-run supply curve to shift: the wage, $\widehat{w}_t$, and the firm's desired margin, $\widehat{\mu}_{it}$. Furthermore, the slope of the supply curve is $\left(\frac{\alpha}{1-\alpha}\right)$. Therefore, provided that

$\alpha < 1$, the supply curve will be upward sloping, since, as output rises (holding technology constant), the marginal product of labour falls as long as $\alpha < 1$.[5]

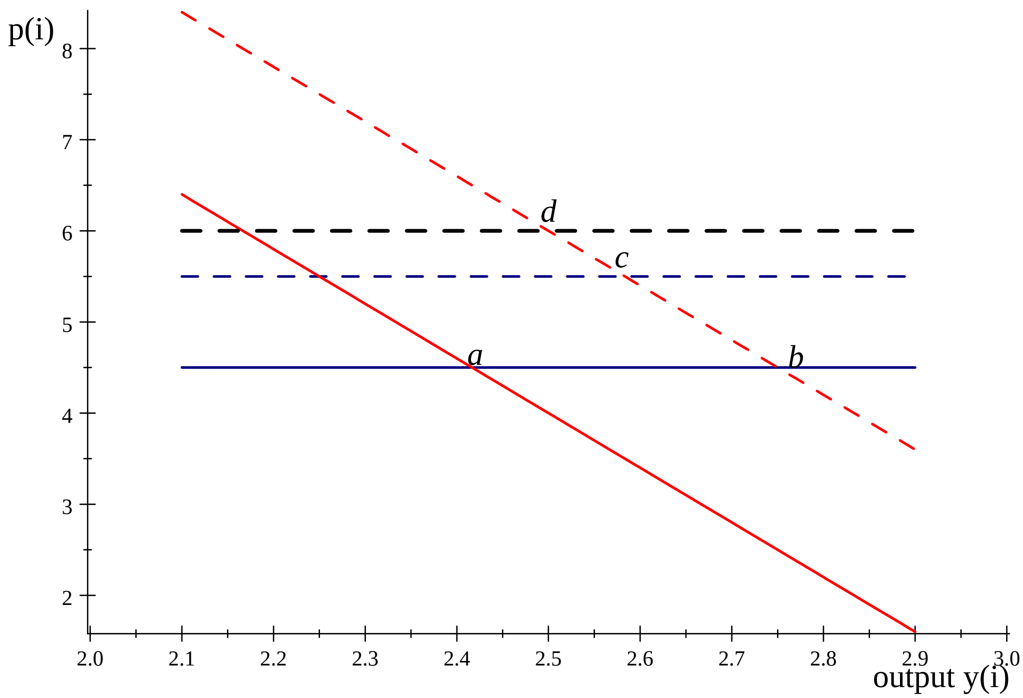
We begin our experiment at point $a$ in Figure 1, which will correspond to our initial equilibrium. Consider a positive shock to demand that pushes the solid red line to the broken red line. In ToTEM, two things will happen. First, an increase in demand will require the firm to hire more labour, which, given equation (1.1), implies a fall in the marginal product of labour and a move up the supply curve to point $b$. In addition, the increase in demand for labour will cause the nominal wage to rise, which, given equation (1.5), will shift the supply curve in and the new equilibrium will be at point $c$. For ToTEM, this is the end of the story, since, by assumption, $\widehat{\mu}_{it} = 0$. For QPM (Figure 2), however, there will be a final shift in of the supply curve to point $d$, where the black and red broken lines intersect, reflecting an increase in the firm's desired markup of price over marginal cost commensurate with the increase in demand, $\widehat{\mu}_{it} > 0$. In addition, the initial shift from $a$ to $b$ represents a pure increase in the quantity supplied, with no corresponding increase in prices, given the assumption in QPM that firm-level marginal cost is independent of output (for a given level of technology). In our simple model, this corresponds to the assumption that $\alpha = 1$.

Of course, these differences do not imply that prices must rise by more or less in the short run in QPM, since this can be adjusted via the model's calibration; rather, it merely illustrates the channels through which a demand shock translates to higher prices in the two models.

To summarize, in both models, a movement in demand affects inflation in two ways: through a direct-demand channel and through a factor-price channel. In ToTEM, since the marginal products of the variable factors are declining, the increase in output implies a decrease in the marginal product and thus an increase in marginal cost. At the same time, the resulting increase in labour demand across firms eventually leads to a higher aggregate wage rate, which also increases real marginal cost. In QPM, the same factor, an increase in the aggregate wage rate, causes a shift in the aggregate supply curve. But it is the output gap, rather than decreasing short-run returns to scale, that provides the direct link between demand in the goods market and price inflation.

[5] In ToTEM, the firm-level marginal-cost curve is positively sloped, because of the assumption of firm-specific capital services (section 2.1). For our purposes, it suffices to simply set $\alpha < 1$.

**Figure 2: Demand/Supply in QPM**




### 1.5.2 *The instrument and transmission of monetary policy*

The inclusion of an endogenous monetary policy reaction function in QPM represented an important innovation relative to previous-generation Bank of Canada models. In contrast to those models, which often treated policy as exogenous, QPM featured an important role for policy in conditioning agents' short- and long-run expectations, and it also made explicit both the instrument and objective of policy through the specification of the rule.[6] While this is also true for ToTEM, the instrument of policy (the variable assumed to be directly under the control of policy) differs across the two models. In QPM, the instrument was assumed to be the adjusted yield spread; i.e., the difference between the 90-day commercial paper rate and a 10-year government bond yield, adjusted for the term premium. At the time, the use of the yield spread was justified on the grounds that it better reflected the stance of policy than did short-term rates. In addition, it provided

[6] Conditional on a functional form for the central bank's loss function.

"a parsimonious way to capture the effects of the full term structure on aggregate spending" (Coletti et al. 1996).

In a sense, these arguments address two separate issues. The first issue is what variable belongs in the reaction function of the monetary authority; i.e., which variable does monetary policy affect most directly? In the current policy framework, this variable is the so-called overnight rate. The overnight rate is the interest rate at which major financial institutions borrow and lend one-day (or "overnight") funds among themselves; the Bank sets a target level for that rate. This target for the overnight rate is often referred to as the Bank's key interest rate or key policy rate. The second issue is which interest rate is most relevant to demand determination in the economy. In QPM, the adjusted yield spread fulfilled both roles. In other words, there was a direct link between the actions of the central bank and, in particular, consumption and investment spending that was independent of the structure of the rest of the model.

In ToTEM, given the assumption of profit and welfare maximization on the part of individuals and firms, the set-up is somewhat more complex. In terms of the instrument of policy, we assume that the Bank can control the 90-day nominal commercial paper rate through its influence on the overnight rate. However, the level of the nominal short-term rate does not, in and of itself, affect real spending. Rather, consumption and investment spending are determined by the entire expected future path of short-term real interest rates (see section 4.3). In this sense, one can think of the demand side of the economy as being influenced by a long-term real interest rate. Furthermore, changes in nominal short-term interest rates influence this long-term real rate only because prices and wages are not fully flexible in the short run. In this sense, the relative influence of monetary policy on nominal versus real variables will depend strongly on the degree of nominal rigidity in general, and the degree of rigidity in labour markets in particular. No such link existed in QPM, and therefore questions relating to changes in average contract length and its implications for the effects of monetary policy could not be considered. Another consequence of the approach followed in ToTEM is that the current stance of policy matters only to the extent that it influences the long-term rate. In other words, policy will exert about the same influence on spending whether it adjusts interest rates by a lot for a short period of time or by a little for a long period of time. Given a central bank's desire not to generate any undue volatility in interest rates, this feature will tend to imply a good deal of inertia in the instrument of policy (see section 3.1.2).

### 1.5.3 *Endogenous labour supply*

In ToTEM, the supply of labour is endogenous both in the long run and in the short run. This makes sense, given that the structure of the dynamic and steady-state models are mutually consistent, with the latter being a special case of the former. In contrast, labour supply was endogenous only in the short run in QPM. Therefore, implicitly, QPM assumed that the same forces that cause workers to increase their labour supply in the short run (such as higher wages) were not at play in the long run. This distinction can be of material importance for certain shocks.

For instance, while the utility function was chosen such that long-run labour supply is invariant to the level of technology in the economy, thereby eliminating a possible trend in average hours stemming from trend productivity growth, certain permanent shocks do have long-run implications for labour supply. As an example, since labour supply is partly determined by the return to working, which is the real, after-tax wage, changes to the direct tax rate on labour income stemming from a change to the government's desired debt-to-GDP ratio will influence steady-state hours worked.

## 1.6 Expectations and the role of information

QPM was the first model used for projections at the Bank of Canada that formally recognized the importance of expectations in determining economic outcomes. In QPM, the ability of the monetary authority to anchor inflation expectations is key to the effectiveness of policy. Relative to its predecessors, QPM represented a leap forward in terms of the role that expectations played in determining the interaction between policy and economic outcomes, but QPM only went partway towards incorporating fully rational expectations. Expectations in QPM were a weighted average of model-consistent expectations, or expectations based on forecasts that use the entire structure of the model, and adaptive expectations, which are based only on extrapolations of past values of the variable in question.

When QPM was built, it was judged that some form of adaptive expectations was still necessary to yield reasonable model dynamics. QPM relied on adaptive expectations largely to help explain the apparent persistence inherent in the macroeconomic data. In terms of inflation, the high degree of persistence that QPM sought to impart was likely due in large part to how analysts viewed the persistence of inflation at the time. First, to a greater extent than today, the degree of inflation persistence was viewed as a structural feature of the economy, instead of a consequence of the monetary policy regime, and partially adaptive expectations helped QPM generate what seemed to be the appropriate degree of inflation persistence. More recently, however, Demers (2003) and Levin and Piger (2004) have shown that, when one properly accounts for breaks in the mean inflation rate, the degree of persistence found in inflation is much smaller, and may vary across policy regimes. Indeed, in the most recent 10 years, the degree of autocorrelation exhibited by core consumer price inflation has declined to almost zero.

A second, and related, factor is the gradual nature of the response of inflation expectations and core inflation to an announced policy change. Historically, achieving a permanent reduction in the rate of inflation involved a non-trivial output loss, as expectations of inflation and thus actual inflation were slow to adjust. Partly adaptive inflation expectations allowed QPM to generate these costly disinflations, but it also implied that inflation expectations should mimic actual inflation even when the central bank is able to credibly commit to returning inflation to the target following shocks. The latter seems doubtful in light of the evidence on inflation expectations since the adoption of official inflation targets in Canada in 1991. Evidence presented in Levin, Natalucci, and Piger (2004) shows that inflation-targeting countries such as Canada, Australia, New Zealand, Sweden, and the United Kingdom have been successful at delinking expectations from realized inflation. This finding is directly at odds with the assumption at the heart of expectations formation in QPM.

While the thinking about inflation persistence has changed significantly since the early 1990s, until recently it remained a challenge to explain why, if expectations are indeed rational, agents make systematic forecast errors during periods of permanent disinflations, as they did in the early 1980s and 1990s. One explanation is tied to the role of central bank credibility. In the recent literature (see Andolfatto, Hendry, and Moran 2002; Erceg and Levin 2003) and in ToTEM, inflation expectations can be sticky if monetary policy is viewed as being less than fully credible. More specifically, when private agents observe an unusual movement in policy interest rates,[7] they are initially unsure whether the movement represents a permanent shift in the inflation target or a temporary deviation from the monetary

[7]More technically, this uncertainty applies when interest rates differ from that predicted by the policy rule in the model. In solving ToTEM, we assume that households and firms know the rule that the Bank of Canada follows.

policy rule. Agents learn about the event's true nature at a speed that depends crucially on the transparency and credibility of the central bank. The less credible the central bank policy, the slower agents learn, the more persistent are inflation forecast errors, and hence the more persistent are actual inflation and output. By introducing the concept of learning about monetary policy, ToTEM generates realistic short-run output losses in the case of less than fully credible disinflations. At the same time, inflation persistence in ToTEM remains reasonably low in shocks when monetary policy credibility is high.

## 1.7 Compromises with ToTEM

Despite our focus on theoretical coherence, the detailed multi-sector nature of ToTEM requires that certain compromises in terms of structure be made for the sake of tractability. For example, prototype versions of ToTEM (Murchison, Rennison, and Zhu 2004) contained a time-to-build constraint for capital that allowed the model to explain the sluggish response of investment to shocks, without the need to appeal to investment adjustment costs. However, when the model was expanded to a multi-sector framework with multiple capital stocks, it became increasingly onerous from a technical perspective to retain the time-to-build framework. Instead, we opted for the costly investment adjustment framework as in CEE. While perhaps less intuitively appealing, this framework allows the model to produce a degree of persistence similar to that seen in the investment data, while at the same time greatly simplifying the model set-up.

On the price and wage side, the original intention was to model nominal rigidities using the staggered-price set-up of Wolman (1999). This model is appealing in the sense that it imposes a finite maximum nominal contract length and allows for the possibility that old contracts are more likely to be renewed than recently signed contracts. Unfortunately, much like the time-to-build model, this set-up makes the size of the model a direct function of the longest contract length, which in the case of wages can be several years.[8] Thus, the version of ToTEM presented here replaces this set-up with the more compact Calvo (1983) approach.

Finally, while we have attempted to adhere as closely as possible to a fully optimizing framework, some compromises have been made that cannot be explicitly justified on the basis of theory. First, rather than using a pure uncovered interest

[8] Other complications arise when firm-specific capital is introduced that are avoided, up to a first-order approximation, using the Calvo (1983) set-up.

rate parity (UIP) condition, which is generally viewed to not hold in the data, we use a hybrid set-up that allows for a lag of the exchange rate to enter what would otherwise be a purely forward-looking specification (see section 2.4). Second, in order to match the response of Canadian GDP in foreign demand shocks, it is necessary to add the rest-of-world output gap to the finished-product export equation (as with QPM), thereby increasing the demand elasticity to something greater than unity (see section 2.4).[9] Finally, as described in section 2.1, the behaviour of hoarded labour (or effort) is imposed on the model, and therefore is not a choice variable for the firm. Effort is included because the model requires that at least one labour-market margin not be compensated in order to be able to match the properties of labour share over the business cycle. However, in a world where variations in effort are costless to the firm and there is no uncertainty, profit maximization would dictate that all output be satisfied using effort. Earlier work focused on adding fixed costs of production to the model, which introduces locally increasing returns to scale. However, calibrating the fixed costs to yield zero profits in the long run (see CEE) was not sufficient to generate procyclical (counter-cyclical) labour productivity (labour's share). Conversely, calibrating fixed costs to replicate the behaviour of labour productivity produced implausibly large values for the fixed costs.

[9] We could have added this variable to any of the equations that make up demand in the model.

# Chapter 2

# Model Description

ToTEM is an open-economy, dynamic stochastic general-equilibrium (DSGE) model with four distinct finished-product sectors as well as a commodity-producing sector. The behaviour of all key variables in ToTEM can be traced to a set of fundamental assumptions about the underlying structure of the Canadian economy, which greatly improves the model's ability to tell coherent, internally consistent stories about how the Canadian economy is—or will be—evolving. The multiple-products approach also allows ToTEM to inform the staff's judgment on a much wider variety of shocks, including relative-price shocks, which was quite difficult with one-good models like QPM that included no role for relative prices.

In ToTEM there are four sets of agents: households, firms, the central bank, and a representative fiscal authority, or government. The first three are modelled as explicitly maximizing an objective, subject to a set of well-defined constraints. For example, firms in the model wish to maximize their profits, but are faced with constraints such as their production technology and the frequency with which they can change their prices. Consumers wish to maximize their well-being or "utility," subject to a budget constraint that limits the rate at which they can accumulate debt. Finally, the central bank in ToTEM wishes to maximize the well-being of consumers by minimizing deviations of inflation from the target and output from potential, as well as the variability of interest rates, while recognizing that the structure of the economy simultaneously constrains its achievement of these joint objectives (Cayen, Corbett, and Perrier 2006).

Fiscal policy is modelled somewhat more traditionally in ToTEM. The government levies direct and indirect taxes and then spends or transfers to consumers the proceeds of these taxes according to a set of rules that are consistent with achieving a pre-specified ratio of debt to GDP over the medium term. The short-run responses of the rules are calibrated to mimic the historical behaviour of fiscal policy in Canada.

Regarding the role played by consumers (or households), ToTEM assumes the existence of two types of consumers, who differ only in their access to asset and credit markets. The first type, labelled "lifetime-income" consumers, face a lifetime budget constraint but can freely borrow or save, so as to reallocate consumption across time. These agents base their consumption decisions on their total expected lifetime income and will thus choose a very smooth consumption path through time when the real interest rate is constant. Higher (lower) real interest rates will cause lifetime consumers to temporarily increase (reduce) their savings, in order to fully exploit the interest rate change. These agents are also assumed to own the domestic companies and are therefore the recipients of any profits.

"Current-income" consumers, in contrast, face a period-by-period budget constraint that equates their current consumption with their disposable income, including government transfers. In addition to not being able to save or dissave, current-income consumers do not own shares in companies and therefore do not receive dividends. The presence of current-income consumers in ToTEM reflects the simple fact that not all households in the economy can access credit markets, as is typically assumed in DSGE models. In terms of model behaviour, the main implication of introducing current-income consumers is that changes to taxes and transfers have larger consumption effects.

Both types of households sell labour to domestic producers and receive the same hourly wage, which they negotiate with the firm. It is important to note that workers are assumed to possess skills that are partially specific to the individual, thereby implying imperfect substitutability across workers. This assumption about the structure of labour markets is important, because it means that workers have some market power in determining their wage. We also assume that workers and firms do not renegotiate the nominal wage every period, but rather do so about once every six quarters, on average. Furthermore, contract renewals are staggered through time, so a constant proportion is renewed each period. The introduction of "sticky" nominal wages will play a crucial role in creating business cycles in ToTEM, while at the same time allowing monetary policy to influence real variables such as GDP in the short run (monetary policy non-neutrality).

In determining the desired real wage of households, the assumption that both consumption and leisure are valued by households implies that, when negotiating their wage, they will consider both their current consumption level and the number of hours they are working. All else being equal, higher consumption or higher labour input will cause households to demand a higher real wage. The former effect occurs because a high consumption level makes leisure relatively more valuable. Thus, the only way to persuade the household to continue working the same number of hours is to offer a higher real wage.

Regarding the set of firms in the model, ToTEM contains producers of four distinct finished products: consumption goods and services, investment goods, government goods, and export goods. Each type of firm combines capital services, labour, commodities, and imports to produce a finished good. In the current version of ToTEM, only the relative import concentration distinguishes these goods in steady state; future versions, however, will allow for differences in the relative intensities of all factor inputs. The production technology for finished goods is characterized by constant elasticity of substitution. Increased capital utilization is possible, but at a cost. In other words, if a firm chooses to use its capital more intensively (by, say, adding an extra shift), the capital stock will effectively age faster, which in turn will reduce its productivity.

In addition to choosing the optimal mix of inputs, firms set a price for their product with the goal of maximizing their expected profits. Under the assumption that the elasticity of demand for any particular firm's product is constant, profit maximization corresponds to choosing a price that is a constant markup over marginal cost.[11] However, as with nominal wages, we assume that prices are costly for the firm to adjust, and therefore firms do so infrequently, and in a staggered fashion. It will therefore not be possible for the firm to maintain a constant markup, except in steady state. Rather, knowing that any price they choose will likely be in effect for several periods, firms will set their nominal price so as to maintain a particular average markup over the duration of the period. Subsequent shocks will then cause variations in the firm's relative price, leading to variations in their sales, with low-price firms capturing greater market share. In addition to the assumption of nominal rigidity, we also allow for an important role to be played by real rigidities at the firm level.[12] Indeed, it is the interaction between a

[11] As in the labour market, the goods market is assumed to be characterized by imperfect competition, which implies that firms have some power to choose a price that differs from the price of their competitors and still remain in business. Marginal cost refers to the cost to the firm of producing one additional unit of output.

[12] By nominal rigidity we refer to mechanisms such as binding contracts that prevent a firm from changing their nominal price to changes in aggregate-demand conditions. Real rigidity refers to any mechanism that reduces the extent to which a firm wishes to change their relative price when given the opportunity to do so.

relatively small degree of nominal rigidity and a large degree of real rigidity that allows us to match the insensitivity of inflation to aggregate real marginal cost, or, for that matter, demand relative to long-run supply, that has been documented in much of the literature that estimates New Keynesian Phillips curves, without having to appeal to counterfactually long nominal contracts.[13] Put another way, pricing decisions in ToTEM are *strategic complements* (Woodford 2003, chapter 3), in the sense that firms have a strong incentive to raise their price when other firms' prices increase, and vice versa.

Imports are treated as inputs to production, rather than as separate final goods. An importing firm buys goods from the foreign economy according to the law of one price, and sells them to manufacturing firms at a price that is also adjusted only periodically. Thus, movements in exchange rates or foreign prices are not fully reflected immediately in the price paid by domestic producers. Furthermore, since the prices of both imported inputs and finished products are sticky, the model includes an element of vertical or supply-chain price staggering, which is crucial in allowing the model to generate realistic exchange rate pass-through to the CPI.

ToTEM also contains a separate commodity-producing sector. Commodities are either used in the production of finished products, purchased directly by households as a separate consumption good, or exported on world markets. The law of one price is assumed to hold for exported commodities, whereas temporary deviations from the law of one price are permitted for commodities that are purchased domestically.

## 2.1 The finished-products sector

We begin by assuming the existence of a continuum of monopolistically competitive firms that produce an intermediate good with a nested CES technology that combines capital, labour, commodities, and imports. We consider here the $i$th firm in the core consumption goods and services sector, but the same functional forms apply for investment goods, government goods, and export goods. We refer to this as the core consumption good since, as discussed in section 2.1.1, total consumption is a composite basket of the core good and commodities. This set-up provides a clean distinction between the core CPI (see Macklem 2001) and the

[13] Much of the literature focuses on U.S. data. ACEL, for instance, estimate that the coefficient on the real marginal cost is about 0.04. Amano and Murchison (2005) and Gagnon and Khan (2005) report similar results for Canada.

total or headline CPI in ToTEM.

For simplicity, it is useful to divide production into three stages. In the first stage, a capital-labour composite is produced with the following CES technology:

$$\mathcal{G}(A_t L_{it}, u_{it} K_{it}) = \left( (\alpha_{c,1})^{\frac{1}{\sigma}} \left( A_t L_{it}^c \right)^{\frac{\sigma-1}{\sigma}} + (1-\alpha_{c,1})^{\frac{1}{\sigma}} \left( u_{it}^c K_{it}^c \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \qquad (2.1)$$

where $u_{it}^c$, $K_{it}^c$, $L_{it}^c$, and $A_t$ are the rate of capital utilization, the level of the capital stock, labour inputs demanded by firm $i$, and the economy-wide level of labour-augmenting technology, respectively. The parameter $\sigma$ governs the elasticity of substitution between capital and labour services; when $\sigma = 1$, equation (2.1) reduces to the more conventional Cobb-Douglas function. $A_t$ evolves according to the first-order autoregressive process:

$$\ln(A_t) = \ln(g) + \ln(A_{t-1}) + \varepsilon_t^A \quad \varepsilon_t^A \sim (0, \sigma_A^2), \qquad (2.2)$$

where $g$ is the gross steady-state growth rate of technology and $L_{it}^c$ is assumed to comprise observed employment, $H_{it}^c$, and unobserved labour effort, $E_{it}^c$:

$$L_{it}^c = H_{it}^c E_{it}^c. \qquad (2.3)$$

Effort may be thought of in one of two ways. It can capture variations in "hoarded" labour and/or it can represent the intensity of work effort in a model of non-hoarded labour. In either case, the result is that firms can adjust, in the short run, the utilization rate of employed workers. It is important to note, however, that the firm in our set-up chooses total labour input, $L_{it}^c$, not the components. In determining the level of effort, we assume an implicit contract between firms and workers, stipulating that variations in work effort can be used in the short run to compensate for the effects of adjustment costs on employment, but, on average, effort equals some normal level ($E^c = 1$).[14] For example, when demand rises (assuming constant technology), firms are required to increase their labour input so as to raise production. But firms can hire new workers only gradually, due to search and hiring costs, and those new workers will be less productive than existing workers owing, for example, to a lack of firm-specific training. Under this agreement, effort will then increase to make up the difference between employed hours and the overall level of labour input required to satisfy demand (causing

[14] Attempts to model effort in a fully optimizing manner have so far proved unsuccessful. This represents an important area for future work.

observed labour productivity to rise). In other words, firms can vary effort such that overall labour input, $H^c_{it}E^c_{it}$, just equals the level that would prevail if there were no labour adjustment costs. This is equivalent to permitting firms to ignore adjustment costs on employment when they choose overall labour input.

While effort clearly benefits the firm, there is no accompanying direct cost to the firm, since we assume that effort is not directly compensated. This assumption is necessary in order to guarantee that labour's share of income is counter-cyclical or, equivalently, that the share of corporate profits in income is procyclical. This aspect of the data, together with the procyclicality of observed labour productivity, is the central motivation for adding a role for effort in the model.

The capital-labour composite is combined with a commodity input to produce the domestic composite:

$$\mathcal{H}\left(\mathcal{G}, COM^c_{it}\right) = \left((\alpha_{c,2})^{\frac{1}{\varrho}}\,\mathcal{G}(A_t L^c_{it}, u^c_{it}K^c_{it})^{\frac{\varrho-1}{\varrho}} + (1-\alpha_{c,2})^{\frac{1}{\varrho}}\left(COM^c_{it}\right)^{\frac{\varrho-1}{\varrho}}\right)^{\frac{\varrho}{\varrho-1}}\ , \quad (2.4)$$

where $COM^c_{it}$ represents commodities or raw materials that are used in the production of the core consumption good. Finally, this domestic composite is combined with imports in the following function:

$$\mathcal{F}\left(\mathcal{H}, M^c_{it}\right) = \left((\alpha_{c,3})^{\frac{1}{\varphi}}\left(A^c_t\mathcal{H}\left(\mathcal{G}, COM^c_{it}\right)\right)^{\frac{\varphi-1}{\varphi}} + (1-\alpha_{c,3})^{\frac{1}{\varphi}}\left(M^c_{it}\right)^{\frac{\varphi-1}{\varphi}}\right)^{\frac{\varphi}{\varphi-1}}\ , \quad (2.5)$$

where $M^c_t$ is the imported consumption good. $A^c_t$ is the consumption-sector specific technology index and evolves according to the unit-root process:

$$\ln(A^c_t) = \ln(A^c_{t-1}) + \varepsilon^{A^c}_t \qquad \varepsilon^{A^c}_t \sim (0, \sigma^2_{A^c}). \quad (2.6)$$

The firm incurs a quadratic cost when it adjusts the level of capital, investment, and hours worked, all of which take the form of a dead-weight loss of the produced good. For the labour adjustment cost, we assume that the productivity of the new workers (as well as those established workers required to train them) will be reduced for two periods, including the hiring period. We also assume that the firm can vary its rate of capital utilization at the cost of foregone output.

When we incorporate quadratic capital, investment, and labour adjustment costs, in addition to convex costs of capital utilization, net output of the core consumption good evolves according to:

$$\begin{aligned} C_{it} &= \mathcal{F}\left(\mathcal{H}, M_{it}^c\right) - \frac{\chi_k}{2}\left(\frac{I_{it}^c}{K_{it}^c} - (\omega + g - 1)\right)^2 K_{it}^c - \frac{\chi_I}{2}\left(\frac{I_{it}^c}{I_{i,t-1}^c} - g\right)^2 I_t \\ &\quad -\rho\left(1 - e^{\chi_u(u_{it}^c - 1)}\right)\frac{P_t^I}{P_t^c}K_{it}^c - \frac{\chi_h}{2}\left(\left(\frac{H_t^c}{H_{t-2}^c}\right) - 1\right)^2 A_t, \end{aligned} \tag{2.7}$$

where $I_{it}^c$ is gross investment by the $i$th firm in the core consumption sector, $\chi_k$, $\chi_I$, and $\chi_h$ determine the size of capital, investment, and labour adjustment costs, and $\chi_u$ and $\rho$ determine the costs of variable capital utilization.[15] The capital stock in period $t+1$ is determined by

$$K_{i,t+1}^c = (1-\omega)K_{it}^c + I, \tag{2.8}$$

where $\omega$ is the quarterly rate of capital depreciation. The $i$th firm's objective is to choose $C_{it}$, $L_{it}^c$ , $K_{i,t+1}^c$, $I_{it}^c$, $u_{it}^c$, $COM_{it}^c$, and $M_{it}^c$ subject to equations (2.7) and (2.8) to maximize the value of the firm:

$$\mathcal{V}_{it} = \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s}\left(P_{is}^c C_{is} - W_s H_{is}^c - P_s^{com} COM_{is}^c - P_s^I I_{is}^c - P_s^m M_{is}^c\right), \tag{2.9}$$

where $P_s^I$ is the price of investment, $P_s^{com}$ is the price of commodities used in production and $W_s$ is the aggregate nominal wage,[16] and $P_s^m$ is the price of the imported good. The stochastic discount factor, $\mathcal{R}_{t,s}$, is defined as

$$\mathcal{R}_{t,s} \equiv \prod_{v=t}^{s-1}\left(\frac{1}{1+R_v}\right) \quad \mathcal{R}_{t,t} \equiv 1 \;\text{ and }\; \mathcal{B} \equiv \frac{g}{1+r} \lessapprox 1.$$

where $R_t$ is the quarterly nominal interest rate and $r$ is the steady-state real quarterly interest rate. $\mathcal{B}$ captures the real, steady-state discount rate, adjusted for trend technology growth.

In what follows, we present only the linearized, aggregated first-order conditions (for the derivation of the non-linear first-order conditions, see Appendix A).

[15] We impose a restriction on the parameter $\rho$ such that, in steady state, the utilization rate is one, which implies that $\rho < 0$. In addition, the adjustment-cost functions for capital, investment, and labour are set up such that adjustment costs are zero in steady state, after accounting for steady-state growth.

[16] The nominal wage is assumed to be equal across sectors in ToTEM.

Labour demand is given by the familiar condition relating the nominal marginal cost of production to the nominal wage divided by the marginal product of labour. However, since both the price level and the technology index contain unit roots, it is more intuitive to deflate the marginal cost of production and the wage by the price level in the consumption sector, and to deflate the marginal product of labour and the real wage by $A_t$. Log linearization of this equation yields:

$$\widehat{\lambda}_t^c = \widehat{w}_t - \widehat{\mathcal{F}}_l\left(\cdot, t\right), \tag{2.10}$$

where $\widehat{\mathcal{F}}_l\left(\cdot, t\right)$ captures the percent deviation of the marginal product of labour input (deflated by $A_t$) from steady state, at time $t$. $\widehat{w}_t$ is the per cent deviation of the nominal wage divided by $A_t P_{t,}^c$ from its steady state, and $\lambda_t^c$ is the Lagrange multiplier associated with the constraint equating supply and demand, so in this context, $\widehat{\lambda}_t^c$ is interpretable as the per cent deviation of real marginal cost (nominal marginal cost deflated by $P_t^c$) from steady state.

As previously mentioned, conditional on the firm's choice of total labour input, the division between employment and effort is determined by the size of the adjustment costs on hours worked, $\chi_h$. Specifically, the optimal condition for employment is found by maximizing the value of the firm with respect to hours, $H_t^c$, which yields:

$$\widehat{H}_t^c \;=\; \left(\frac{1}{1+\mathcal{B}^2}\right)\widehat{H}_{t-2}^c + \left(\frac{\mathcal{B}^2}{1+\mathcal{B}^2}\right)\widehat{H}_{t+2}^c - \frac{s}{\chi_h\left(1+\mathcal{B}^2\right)}\left(\widehat{w}_t - \widehat{\lambda}_t^c - \widehat{\mathcal{F}}_h\left(\cdot, t\right)\right), \tag{2.11}$$

where $s = \left(\frac{\overline{w_t H_t^c}}{A_t}\right)$ is proportional to labour's share of income in steady state. Demand for employment is driven by a forward-looking error-correction equation that depends positively on hours two periods ahead and previous, and on the cointegrating relationship between the wage (deflated by marginal cost) and the marginal (gross) product of hours. Effort is then proportional to the difference between the marginal product of hours and the real wage:

$$\widehat{E}_t^c = \widehat{\mathcal{F}}_h\left(\cdot, t\right) - \left(\widehat{w}_t - \widehat{\lambda}_t^c\right). \tag{2.12}$$

Note that, as $\chi_h \longrightarrow 0$, the cointegrating term in equation (2.11) will always hold, and therefore effort, as determined by equation (2.12), will always be zero and hours and total labour input will be equal. Otherwise, if $\chi_h > 0$ , hours will adjust only gradually to a shock, and effort will adjust such that (2.10) holds.

The linearized first-order conditions for import and commodity inputs are

$$\widehat{\lambda}_t^c = \widehat{p}_t^m - \widehat{\mathcal{F}}_m(\cdot, t), \tag{2.13}$$

and

$$\widehat{\lambda}_t^c = \widehat{p}_t^{com} - \widehat{\mathcal{F}}_{com}(\cdot, t), \tag{2.14}$$

respectively, where each price is again deflated by $P_t^c$. Analogous to the labour input, firms choose the quantity of imports and commodities such that the marginal product of that input equals its price deflated by marginal cost.

We turn next to the firm's decisions regarding capital. We start with the shadow value of capital, $q_{it}^c$, or the discounted contribution of capital to future dividends, which is given by:

$$\widehat{q}_t^c = \frac{1}{1+r}\mathbf{E}_t \left( \begin{array}{c} (r+\omega)\left(\widehat{\lambda}_{t+1}^c + \widehat{\mathcal{F}}_k(\cdot, t+1) - \widehat{u}_{t+1}^c\right) + \\ \chi_K(\omega+g-1)^2\left(\widehat{i}_{t+1}^c - \widehat{k}_{t+1}^c\right) + (1-\omega)\,\widehat{q}_{t+1}^c - \widehat{r}_t \end{array} \right), \tag{2.15}$$

where $\widehat{r}_t$ captures variations in the ex ante real interest rate in the consumption sector, relative to steady state. $\widehat{q}_t^c$ captures the present discounted net value to the firm, in real terms, of an additional unit of installed capital. If we ignore capital adjustment costs ($\chi_K = 0$) and capital utilization for the moment, and solve equation (2.15) forward, we obtain:

$$\begin{aligned} \widehat{q}_t^c \quad = \quad & \frac{r+\omega}{1+r}\mathbf{E}_t \sum_{s=t}^{\infty} \left(\frac{1-\omega}{1+r}\right)^{s-t} \left(\widehat{\lambda}_{s+1}^c + \widehat{\mathcal{F}}_k(\cdot, s+1)\right) \\ & -\frac{1}{1+r}\mathbf{E}_t \sum_{s=t}^{\infty} \left(\frac{1-\omega}{1+r}\right)^{s-t} \widehat{r}_s. \end{aligned} \tag{2.16}$$

The first line of expression (2.16) captures the present discounted value of the marginal benefits (or shadow values) from period $t+1$ into the infinite future accruing to the firm from a marginal change to the capital stock in period $t$, assuming a constant real interest rate. Note that, in discounting the future, the depreciation rate $\omega$ is also factored into the calculation of the present value. If one unit of capital is installed in period $t$, $(1-\omega)^{s-t}$ will remain in period $s$. The term on the second line captures expected variations in the real interest rate, which cause variations in the effective discount rate. When we add back the adjustment costs on capital and variable capital utilization, the calculation of the present value is somewhat more complicated. But the intuition remains the same: $q$ captures the net marginal benefit to the firm of an additional unit of capital. Thus, if $\widehat{q}^c_{it} > \widehat{p}^I_t$, then the firm has an incentive to raise its investment. The assumption of adjustment costs on changes in investment means that investment responds only gradually to changes in $\widehat{q}^c_{it}$ relative to $\widehat{p}^I_t$, as shown by the linearized investment equation:

$$\widehat{i}^c_t = \Im\left(\widehat{i}^c_{t-1} - \widehat{g}_t\right) + \mathcal{B}\Im\left(\widehat{i}^c_{t+1} + \widehat{g}_{t+1}\right) + \left(1 - \Im\left(1 + \mathcal{B}\right)\right)\widehat{k}^c_t + \Im\Gamma\left(\widehat{q}^c_t - \widehat{p}^I_t\right), \tag{2.17}$$

with $\Im = \frac{(1+r)g^2\chi_I}{g^3\chi_I + (1+r)(g^2\chi_I + \chi_K(\omega+g-1))} \approx 0.5$, $\Gamma = \frac{p^I}{\chi_I\lambda g^2}$, and $g$ again the gross steady-state real growth rate of the economy. Like hours, log deviations of investment are driven by a forward-looking error-correction equation. In the absence of investment adjustment costs ($\chi_I = 0$), the level of investment will adjust instantaneously to ensure that deviations of the shadow value of capital, $\widehat{q}_t$, are always equal to deviations of the price of investment, $\widehat{p}^I_t$. With adjustment costs, firms will optimally trade off the costs of having a suboptimal level of investment $\left(\widehat{q}^c_{it} \neq \widehat{p}^I_{it}\right)$ with the costs of adjusting their level of investment. A lag of investment will therefore appear in equation (2.17), which helps the model generate a hump-shaped investment response to movements in interest rates, consistent with the empirical evidence for Canada presented in Murchison, Rennison, and Zhu (2004).

The linearized first-order condition for capital utilization is then given by

$$\widehat{u}^c_t = \frac{1}{\chi_u}\left(\widehat{\mathcal{F}}_u(\cdot, t) - \widehat{p}^I_t - \widehat{k}^c_t\right). \tag{2.18}$$

Not surprisingly, a high marginal product of capital utilization will encourage higher capital utilization, while a high relative price of investment or capital stock will tend to reduce capital utilization. These latter effects are present due to the assumption that the costs of capital utilization are measured in units of capital multiplied by the relative price of investment. The intuition is that a high rate of capital utilization will be more costly to a firm when the stock of capital is high, or when the replacement cost of capital is high ($\widehat{p}^I_t$). Thus, while we do not explicitly model capital utilization as affecting the rate of depreciation for capital ($\omega$), our adjustment-cost model captures the essence of this effect.

We turn next to the firm's pricing decision. As mentioned, we assume a continuum of monopolistically competitive firms that each produce a differentiated consumption good and charge a price for their good that maximizes expected profits. Thus, the representative firm, $i$, $i \in [0,1]$, will produce $C_i$ and receive price $P_i^c$ in return. Aggregate core consumption, $C_t$, and its corresponding deflator, $P_t^c$, are given as:

$$C_t \equiv \left( \int_0^1 C_{it}^{\frac{\epsilon_t^c - 1}{\epsilon_t^c}} \, di \right)^{\frac{\epsilon_t^c}{\epsilon_t^c - 1}}, \tag{2.19}$$

$$P_t^c = \left( \int_0^1 (P_{it}^c)^{1-\epsilon_t^c} \, di \right)^{\frac{1}{1-\epsilon_t^c}}. \tag{2.20}$$

Note that we treat the elasticity of substitution between finished consumption goods ($\epsilon_t^c$) as stochastic, as in Smets and Wouters (2005). In addition, we assume the following process:

$$\ln(\epsilon_t^c) = (1 - \rho_\epsilon^c) \ln(\epsilon) + \rho_\epsilon^c \ln(\epsilon_{t-1}^c) + \varepsilon_t^{\epsilon_c} \qquad \varepsilon_t^{\epsilon_c} \sim (0, \sigma_{\epsilon_c}^2). \tag{2.21}$$

The degree of market power is assumed to be equal across sectors in ToTEM, with the notable exception of the commodity sector, and therefore the steady-state elasticity of substitution is not indexed to the consumption sector ($\epsilon$). Cost minimization in the production of a unit of $C$ by the aggregator implies that firm $i$ faces the demand schedule

$$C_{it} = \left( \frac{P_{it}^c}{P_t^c} \right)^{-\epsilon_t^c} C_t \tag{2.22}$$

for its product. In addition, we assume that firms not chosen to reoptimize can nevertheless update their price according to a geometric average of lagged CPIX inflation and the current-period expectation of the inflation target (see section 2.6 for details on the target):

$$\Pi_{s,t}^c \equiv \left( \frac{P_{s-1}^c}{P_{t-1}^c} \right)^{\gamma_c} \mathbf{E}_t \left( \prod_{j=t}^{s-1} (1 + \overline{\pi}_j) \right)^{1-\gamma_c} \quad s > t. \tag{2.23}$$

The first-order condition for a reoptimizing firm is given by

$$P^c_{i,t} = \mathbf{E}_t \frac{\sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta_c^{s-t}\lambda_{is}C_{is}\epsilon^c_s}{\sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta_c^{s-t}\left(\Pi^c_{s,t}\right)C_{is}\left(\epsilon^c_s - 1\right)}. \tag{2.24}$$

The aggregate price level for the consumption sector at time $t$, using equation (2.20), is given by

$$P^c_t = \left(\theta_c\left(P^c_{t-1}\Pi^c_{t-1,t-2}\right)^{1-\epsilon^c_t} + (1-\theta_c)\left(P^c_{it}\right)^{1-\epsilon^c_t}\right)^{\frac{1}{1-\epsilon^c_t}}. \tag{2.25}$$

Finally, the log-linearized Phillips curve for CPIX inflation is given by

$$\begin{aligned}\widehat{\pi}^c_t \;=\; & \frac{\gamma_c}{1+\mathcal{B}\gamma}\left(\widehat{\pi}^c_{t-1} - \Delta\mathbf{E}_t\overline{\pi}_t\right) + \frac{\mathcal{B}}{1+\mathcal{B}\gamma_c}\mathbf{E}_t\widehat{\pi}^c_{t+1} \\ & +\zeta\frac{(1-\theta_c)(1-\theta_c\mathcal{B})}{\theta_c(1+\mathcal{B}\gamma_c)}\widehat{\lambda}^c_t + \varepsilon^p_t, \end{aligned} \tag{2.26}$$

where $\widehat{\pi}^c_t$ will correspond to the difference between actual core consumer price inflation and the current perception of firms regarding the central bank's inflation target, $\mathbf{E}_t\overline{\pi}_t$. Therefore, while actual inflation inherits the unit root assumed for the central bank's inflation target, the difference between the two is stationary. The parameter $\zeta$ is governed by our assumption regarding the structure of the capital market. If capital is assumed to be freely tradable among firms in each period, then $\zeta = 1$. If, on the other hand, we make the arguably more plausible assumption that capital is firm owned, specific to that firm, and costly to adjust, then $0 < \zeta < 1$ (Woodford 2005). $\zeta$ is a highly non-linear function of several of the model's key structural parameters and must be solved for numerically (see Appendix A):

$$\zeta = \zeta\left(\mathcal{B}, \theta_c, \epsilon, \sigma, \varrho, \varphi, \omega, \chi_K, \chi_I, \chi_u\right) : \zeta \in (0, 1]. \tag{2.27}$$

In addition, $\zeta$ will depend on the relative importance of capital in the production process, or, more specifically, the elasticity of the gross output of the consumption good with respect to capital. $\zeta$ is decreasing in the degree of competition, $\epsilon$; therefore, in aggregate, inflation will be less sensitive to movements in aggregate marginal cost when markets are highly competitive. As a result there is greater real rigidity or strategic complementarity across pricing decisions. The intuition for this result is that, as the degree of competition increases, firms will be less inclined to increase their own relative price in response to an increase in aggregate costs, since such an increase will lead to a reduction in demand for their product and thus a fall in their own level of marginal cost[17]:

$$\widehat{\lambda}_{it}^{c} - \widehat{\lambda}_{t}^{c} = -\phi_1 \left( \epsilon \widehat{p}_{it}^{c} + \left( \widehat{k}_{it}^{c} - \widehat{k}_{t}^{c} \right) \right), \tag{2.28}$$

with $\phi_1 > 0$ (see Appendix A for the definition of $\phi_1$). Equation (2.28) shows that the wedge between the $i$th firm's real marginal cost and aggregate marginal cost for a given sector is a function of their relative price, $\widehat{p}_{it}^{c}$, and their relative capital stock. Focusing on the former, as markets grow more competitive, the wedge becomes more sensitive to the firm's relative price, thereby increasing the degree of real rigidity in the model.

The assumption that capital is owned by the firm, rather than by the consumer, is often referred to in the literature as firm-specific capital (Woodford 2003). However, in ToTEM, a more accurate name would be firm-specific capital *services*, since the degree to which capital is utilized is also a decision made by the firm. Hence, as shown in equation (2.27), the slope of the firm-level marginal-cost curve also depends on $\chi_u$, which is the parameter governing the costs of varying one's rate of capital utilization. To see why this parameter matters, consider a situation where capital is firm-specific and predetermined at time $t$, but capital utilization is costless to adjust. In this situation, current capital services would remain costless to adjust even if the stock of capital was not. Hence, this is equivalent to the assumption of a rental market for capital, in the sense that firms can costlessly raise (or lower) the level of every input (holding factor prices constant, which are taken as given at the firm level), and hence their marginal cost is independent of their production level ($\zeta = 0$ if $\chi_u = 0$).

What, then, does equation (2.26) tell us about the behaviour of core CPI inflation in ToTEM? To answer this, it is first useful to solve the equation forward, thereby eliminating the $\mathbf{E}_t \widehat{\pi}_{t+1}^{c}$ term[18]:

$$\widehat{\pi}_{t}^{c} = \gamma_c \widehat{\pi}_{t-1}^{c} + \mathbf{E}_t \frac{\zeta \left(1-\theta_c\right)\left(1-\theta_c \mathcal{B}\right)}{\theta_c} \sum\nolimits_{i=0}^{\infty} \left(\mathcal{B}^i\right) \widehat{\lambda}_{t+i}^{c} + \frac{1+\mathcal{B}\gamma_c}{1-\mathcal{B}\rho_{\epsilon}^{c}} \varepsilon_t^{p}. \tag{2.29}$$

Equation (2.29) indicates that the current inflation rate is a function of lagged

[17]See the discussion of the risk-premium shock in chapter 4 for an example of this effect.

[18]We set $\Delta E_t \overline{\pi}_t = 0$ for convenience. Also, note that we exploited the fact that $E_t \mathcal{B}^j \widehat{\pi}_{t+j}^{c} \rightarrow 0$ as $j$ tends to infinity, which simply states that inflation will always return to the target (in expectation), given sufficient time.

inflation (provided $\gamma_c > 0$) and the discounted sum of all future real marginal-cost deviations (as well as a shock term). In other words, it is not just the current value of marginal cost that is relevant to price determination, but the current and all future (expected) marginal costs. Furthermore, since $\mathcal{B}$ =0.993, given our assumptions regarding the steady-state growth rate of output and the steady-state real interest rate, one can essentially disregard the role of discounting. Also, as discussed in chapter 3, $\gamma_c$ is calibrated to be 0.18, meaning that lagged inflation plays a small role in determining current inflation. Consequently, while prices in ToTEM are sticky, the rate of core CPI inflation is actually quite flexible, implying that monetary policy can exert an important influence on current inflation even if it does not exert a large influence over the current value of real marginal cost. Indeed, by adjusting policy to influence expectations of future marginal costs, policy can have a direct effect on inflation in the current period. This is what is typically referred to as the *expectations channel*, and it is very important in ToTEM.

Since the inputs to production are endogenous choice variables from the firm's perspective, expression (2.10) is perhaps not the most useful way to express marginal cost.[19] Ideally, we would like to express it exclusively in terms of those variables that the firm takes as exogenous to their decisions (i.e., the relative prices of each input to production), but maintain the assumption of profit-maximizing behaviour. Unfortunately, incorporating the cost of capital into the marginal-cost expressions in the presence of adjustment costs on capital would make it an extremely complicated, dynamic expression, since capital accumulation affects not just current, but also future, profits. We can, however, express it in terms of real consumer wages deflated by technology, the relative prices of investment, commodities, and imports, as well as the level of capital utilization in that sector.

Given the market structure detailed above, the sector-wide real marginal-cost gap can be written as,

$$\widehat{\lambda}_t^c = \omega_1 \widehat{w}_t + \omega_2 \widehat{p}_t^{com} + \omega_3 \widehat{p}_t^{m} + \omega_4 \widehat{p}_t^{I} + \omega_5 \widehat{u}_t^{c}. \tag{2.30}$$

For the current calibration of ToTEM, the weights $\omega_1, \omega_2, ..., \omega_5$ for the core consumption-product sector are provided in Table 1.

[19] Any one of (2.10),(2.13), or (2.14) could be used to express marginal cost.

**Table 1: Weights Used in Marginal Cost for the Core CPI (eq. (2.30))**

| Coefficient | Value |
|---|---|
| $\omega_1$ | 0.61 |
| $\omega_2$ | 0.10 |
| $\omega_3$ | 0.28 |
| $\omega_4$ | 0.27 |
| $\omega_5$ | 1.36 |

### 2.1.1 *The CPI and core CPI in ToTEM*

To allow for a distinction between the core CPI and the more general definition of consumer prices (the CPI, which includes the consumption of energy as well as the effects of indirect taxes), we define a composite consumption basket, $C_t^{tot}$, that is a CES aggregate of the core finished product, $C_t$, from the previous section and commodities, $C_t^{com}$:

$$C_t^{tot} = \left( \alpha_c^{\frac{1}{\psi}} \left( C_t \right)^{\frac{\psi-1}{\psi}} + \left(1-\alpha_c\right)^{\frac{1}{\psi}} \left( C_t^{com} \right)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}}, \tag{2.31}$$

where $\psi$ governs the elasticity of substitution between the two consumption goods.[20] The associated price index is given by:

$$P_{c,t}^{tot} = \left(1+\tau_{c,t}\right) \left( \alpha_c \left( P_t^c \right)^{1-\psi} + \left(1-\alpha_c\right) \left( P_t^{com} \right)^{1-\psi} \right)^{\frac{1}{1-\psi}}, \tag{2.32}$$

where $P_t^c$ is the price of the finished product, and corresponds to a conceptual definition of the core CPI in that it excludes consumer goods the prices of which are directly linked to commodity prices, as well as the effects of movements in the indirect tax rate, $\tau_{c,t}$. The demand for commodities relative to the core consumption product is given by:

$$\frac{C_t^{com}}{C_t} = \left( \frac{1-\alpha_c}{\alpha_c} \right) \left( \frac{P_t^{com}}{P_t^c} \right)^{-\psi}. \tag{2.33}$$

[20] The inclusion of commodities is intended to capture gasoline, home-heating oil, and natural gas, which are essentially raw materials that are directly consumed by households.

## 2.2 Import sector

We assume the existence of a continuum of intermediate imported goods, $M_{jt}$, $j \in [0,1]$, that are bundled into an aggregate import, $M_t$, by the aggregator and sold to final-goods producers,

$$M_t \equiv \left( \int_0^1 (M_{jt})^{\frac{\epsilon_t^m - 1}{\epsilon_t^m}} dj \right)^{\frac{\epsilon_t^m}{\epsilon_t^m - 1}}. \tag{2.34}$$

Demand by the aggregator for the differentiated goods is given by the familiar cost-minimizing demand functions,

$$M_{jt} = \left( \frac{P_{jt}^m}{P_t^m} \right)^{-\epsilon_t^m} M_t, \tag{2.35}$$

where $P_t^m$ is the aggregate import-price deflator, given as

$$P_t^m = \left( \int_0^1 \left(P_{jt}^m\right)^{1-\epsilon_t^m} dj \right)^{\frac{1}{1-\epsilon_t^m}}. \tag{2.36}$$

We follow Smets and Wouters (2002) in assuming that the price of the imported good is temporarily rigid in the currency of the importing country. Consequently, exchange rate pass-through to import prices is partial in the short run and complete in the long run. Exchange rate fluctuations are absorbed by the importers' profit margins in the short run, since they purchase goods according to the law of one price. Importers therefore take into consideration the future path of foreign prices and the nominal exchange rate when deciding on their time $t$ price. As for the specific form of the nominal rigidity, we again follow the set-up of Calvo (1983), but allow for partial indexation to lagged import-price inflation. Thus, the first-order condition for a firm chosen to reoptimize its price is quite similar to equation (2.24) from the previous section, except for the definition of real marginal cost, which is now the foreign price of the import multiplied by the nominal exchange rate, deflated by the Canadian-dollar price of imports.

When log-linearized about its steady state, the rate of inflation for imported intermediate-good prices is given by

$$\begin{aligned}\widehat{\pi}_t^m &= \frac{\gamma_m}{1+\mathcal{B}\gamma_m}\left(\widehat{\pi}_{t-1}^m - \Delta\mathbf{E}_t\overline{\pi}_t\right) + \frac{\mathcal{B}}{1+\mathcal{B}\gamma_m}\mathbf{E}_t\widehat{\pi}_{t+1}^m \\ &\quad + \frac{(1-\theta_m)(1-\mathcal{B}\theta_m)}{\theta_m(1+\mathcal{B}\gamma)}\widehat{\lambda}_t^m + \varepsilon_t^{pm}, \qquad (2.37)\end{aligned}$$

where $\widehat{\pi}_t^m$ is the difference between import-price inflation and firms' current perception of the central bank's inflation target.

## 2.3 Commodity sector

We assume that a representative, perfectly competitive domestic firm produces commodities and either sells them to a distributor or exports them to the rest of the world. In either case, the firm receives the rest-of-world price of commodities adjusted for by the nominal Canada/rest-of-world exchange rate for its product, $e_t P^*_{com,t}$. The commodity is produced by combining capital services, labour, and land in a nested CES production function. In the first stage, a capital-labour composite is produced using the following CES function:

$$\mathcal{G}(\cdot,t) = \left((\alpha_{com,1})^{\frac{1}{\upsilon}}\left(A_t H_t^{com}\right)^{\frac{\upsilon-1}{\upsilon}} + (\alpha_{com,2})^{\frac{1}{\upsilon}}\left(u_t^{com}K_t^{com}\right)^{\frac{\upsilon-1}{\upsilon}} + A_t\mathcal{F}\right)^{\frac{\upsilon}{\upsilon-1}}, \qquad (2.38)$$

where $\mathcal{F}$ is a fixed factor of production and may be thought of as land. The commodity-producing firm chooses the quantity to produce, $COM_t$, which corresponds to gross output, $\mathcal{G}(\cdot,t)$, less the adjustment costs on each input, in order to maximize the value of the firm, or the discounted flow of profits:

$$\mathcal{V}_t = \mathbf{E}_t\sum_{s=t}^{\infty}\mathcal{R}_{t,s}\left(COM_s e_s P^*_{com,s} - W_s H_s^{com} - P_s^I I_s^{com}\right),$$

The first-order conditions for $H_t^{com}$, $K_{t+1}^{com}$, $I_t^{com}$, and $u_t^{com}$ take the same form as their analogues in the finished-products sector. The sole difference is that, given the presence of perfect competition, we can replace marginal cost, $\lambda_t$, with the domestic price of commodities, $e_t P^*_{com,t}$.

As in the import sector, we assume the existence of a continuum of imperfectly competitive commodity distributors who purchase raw commodities from the Canadian commodity producer at price $e_t P^*_{com,t}$. These commodities are then repackaged and sold to a perfectly competitive aggregator. The aggregator sets a price, $P^{com}_t$, and sells either to final-goods producers or directly to the consumer. Since the set-up is identical to that of the import distributor, the inflation equation for the determination of $P^{com}_t$ is analogous to equation (2.37), except that the marginal cost of production is $e_t P^*_{com,t}/P^{com}_t$. The introduction of a distribution sector effectively limits the degree to which exchange rate or world commodity-price movements pass through to the prices of finished products such as the CPI (or CPIX) in the short run.

Commodities produced in Canada must either be used for domestic production of finished goods, consumed directly by households, or exported:

$$COM_t = C^{com}_t + COM^c_t + COM^I_t + COM^g_t + COM^{ncx}_t + X_{com,t}. \tag{2.39}$$

## 2.4 Households

ToTEM assumes the existence of two types of consumers, who differ according to their access to asset and credit markets. The first type, who are labelled "lifetime-income" consumers, face a lifetime budget constraint but can freely borrow or save to reallocate consumption across time. In other words, these agents base their consumption decision on the total income they expect to earn over their lifetime. These agents are also assumed to own the domestic companies and therefore are the recipients of any profits.

"Current-income" consumers, by contrast, face a period-by-period budget constraint that equates their current consumption with their disposable income, including government transfers. In addition to not being able to save or dissave, current-income consumers do not own shares in companies and therefore do not receive dividends. We also assume for simplicity that current-income consumers receive the aggregate average wage among lifetime-income households in each period.

Aggregate consumption is the sum of consumption of lifetime-income households, $C^l$, and current-income households, $C^c$:

$$C^{tot}_t = C^c_t + C^l_t. \tag{2.40}$$

A continuum of lifetime-income households indexed by $h$, $h \in [0,1]$, purchases finished goods and consumes leisure to maximize their lifetime utility. Following Erceg, Henderson, and Levin (2000), each household is assumed to supply differentiated labour services to the intermediate-goods sector. Household labour services are purchased by an aggregator and bundled into composite labour according to the Dixit-Stiglitz aggregation function:

$$H_t \equiv \left( \int_0^1 (N_{ht})^{\frac{\epsilon_{w,t}-1}{\epsilon_{w,t}}} \, dh \right)^{\frac{\epsilon_{w,t}}{\epsilon_{w,t}-1}}. \tag{2.41}$$

Similarly, the aggregate nominal-wage index is given as

$$W_t = \left( \int_0^1 (W_{ht})^{1-\epsilon_{w,t}} \, dh \right)^{\frac{1}{1-\epsilon_{w,t}}}, \tag{2.42}$$

where we assume a stochastic markup for wages where $\epsilon_{w,t} \sim (\epsilon, \sigma^2_{\epsilon_w})$. The aggregator purchases differentiated labour services to minimize costs. Thus, the demand for labour services from individual $h$ is given as

$$N_{ht} = \left( \frac{W_{ht}}{W_t} \right)^{-\epsilon_{w,t}} H_t. \tag{2.43}$$

Thus, consumer $h$ who signs a contract in period $t$ will have a nominal wage equal to $W_{h,t}\Pi^w_{s,t}$ in period $s$ with

$$\Pi^w_{s,t} \equiv \left( \frac{W_{s-1}}{W_{s-t-1}} \right)^{\gamma_w} \mathbf{E}_t \left( \overline{\Pi}^w_{s,t} \right)^{1-\gamma_w}, \tag{2.44}$$

provided they are not selected to renegotiate their wage during that interval of time.

The instantaneous utility function for the representative lifetime-income household is given as[21]

$$\mathbb{U}_t = \frac{\mu}{\mu-1} (C^l_t - \xi C^l_{t-1})^{\frac{\mu-1}{\mu}} \exp\left( \frac{\eta(1-\mu)}{\mu(1+\eta)} \int_0^1 (E_t N_{ht})^{\frac{1+\eta}{\eta}} \, dh \right), \tag{2.45}$$

[21]Equation (2.46) is non-standard primarily in the sense that consumption and leisure are not additively separable (see King, Plosser, and Rebelo 1988). Consequently, the marginal utility of consumption (leisure) will depend on labour (consumption).

where the $h$ subscript indexes the level of labour input across individuals in each household. Household consumption depends negatively on lagged consumption according to the parameter $\xi$. Thus, we assume that individuals enjoy high consumption in and of itself (provided $\xi < 1$), but that they also derive utility from high consumption today, relative to the previous period (internal habit formation). Alternatively, we can rewrite the term $C_t^l - \xi C_{t-1}^l$ as $(1-\xi) C_t^l + \xi \Delta C_t^l$, which states that agents get utility from both the level and the change in their consumption, with the relative weights determined by $\xi$. In this sense, the utility function with habits will imply a role for smoothing of both the level and growth rate of consumption.

The lifetime-income households maximize lifetime utility according to

$$\mathbf{E}_t \sum_{s=t}^{\infty} \boldsymbol{\beta}_{t,s} \mathbb{U}_{hs}^l \quad \text{with } \boldsymbol{\beta}_{t,s} \equiv \prod_{v=t}^{s-1} \beta_v \qquad \boldsymbol{\beta}_{t,t} \equiv 1, \tag{2.46}$$

subject to the dynamic budget constraint

$$\begin{aligned} & P_{c,t}^{tot} C_t^l + \frac{B_{t+1}}{1+R_t} + \frac{e_t B_{t+1}^*}{(1+R_t^*)(1+\kappa_t)} \\ = \quad & B_t + e_t B_t^* + (1-\varkappa)\left((1-\tau_{w,t})\int_0^1 N_{ht} W_{ht} dh + TF_t\right) + \sum_k \int_0^1 \Delta_{i,k,t} di, \end{aligned} \tag{2.47}$$

where $B_t^*$ and $B_t$ are, respectively, the value of foreign (domestic) currency-denominated bonds held at time $t$ and $e_t$ is the Canadian-dollar price of unit of foreign exchange. The term $\varkappa$ is the share of current-income consumers in the economy. It is worth highlighting that labour effort is not compensated by assumption, and thus does appear in the household's budget constraint. The term $\tau_{w,t}$ is the tax rate on labour income, $TF_t$ is government transfers, and $\Delta_{i,k,t}$ represents profits paid by firm $i$ in production-sector $k$ (consumption, investment products, etc.) to lifetime-income households.

The term $\kappa_t$ is interpreted as the country-specific risk premium and is assumed to have both a deterministic and stochastic component. More specifically, the risk premium will move with the home country's net foreign indebtedness as a share of nominal GDP, as in Schmitt-Grohé and Uribe (2003), thereby ensuring a stationary dynamic path for the net-foreign-asset-to-GDP ratio about its steady-state value. In addition, we assume that the risk premium is subject to a shock process that represents unforecastable changes in investors' preferences:

$$\kappa_t = \varsigma \left[ \exp\left( \frac{e_t B_t^*}{P_t Y_t} \right) - 1 \right] + \varepsilon_t^\kappa, \tag{2.48}$$

$$\varepsilon_t^\kappa = \rho_\kappa \varepsilon_{t-1}^\kappa + \nu_{\kappa,t} \quad \nu_{\kappa,t} \sim (0, \sigma_\kappa^2). \tag{2.49}$$

Maximizing (2.46) with respect to $C_t^l, W_{ht}, N_{ht}, B_{t+1}^*$, and $B_{t+1}$ subject to (2.43) and (2.47), and log-linearizing about steady state, yields the household's first-order conditions. The conventional uncovered interest parity (UIP) condition is obtained by combining the first-order conditions for domestic and foreign bond holdings:

$$\widehat{e}_t = \mathbf{E}_t \widehat{e}_{t+1} + \widehat{R}_t^* - \widehat{R}_t + \widehat{\kappa}_t.$$

To better match the business-cycle properties of the nominal and real exchange rate, we specify a hybrid version of the UIP condition where

$$\widehat{e}_t = \varpi \widehat{e}_{t-1} + (1-\varpi)\mathbf{E}_t \widehat{e}_{t+1} + \widehat{R}_t^* - \widehat{R}_t + \widehat{\kappa}_t \quad \varpi \in [0,1). \tag{2.50}$$

Log-linearizing the consumption first-order condition yields the consumption Euler equation for forward-looking households:

$$\begin{aligned} \widehat{c}_t^l \;=\; & \widehat{c}_{t+1}^l + \Lambda\xi \left( g^{\frac{1}{\mu}} \left( \widehat{c}_{t-1}^l + \Delta\widehat{g}_{t+1} \right) - g\beta \left( \widehat{c}_{t+2}^l + \Delta\widehat{g}_{t+2} \right) \right) \\ & -\Lambda (g-\xi)(1-\mu) \overline{L}^{1+1/\eta} \left( g^{1/\mu} \Delta\widehat{L}_{t+1} - \xi\beta\Delta\widehat{L}_{t+2} \right) \\ & +\Lambda (g-\xi) \left( g^{1/\mu} - \xi\beta \right) \left( \widehat{g}_{t+1} - \frac{\mu}{1+r} \widehat{r}_{c,t} \right) + \varepsilon_t^c, \end{aligned} \tag{2.51}$$

where $\Lambda = \left( g\xi\beta + g^{1/\mu+1} + \xi^2\beta \right)^{-1}$; $\varepsilon_t^c = -\Lambda (g-\xi) \left( g^{1/\mu} - \xi\beta \right) \mu \widehat{\boldsymbol{\beta}}_{t,t+1}$, and is positive; and $\widehat{r}_{c,t} = \widehat{R}_t - \pi_{c,t+1}^{tot}$. $\widehat{c}_t^l$ denotes per cent deviations of lifetime-income consumption relative to the technology index, $A_t$. Thus, in general, consumption at time $t$ relative to steady state depends positively on a lag and a lead of consumption, and negatively on consumption led two periods. Current consumption depends negatively on the expected real interest rate and, because of our choice of a non-additive utility function, positively on current labour input. Specifically, with $\mu < 1$, consumption will increase in response to a rise in labour input, everything else held constant. Essentially, when agents are working a great deal (and leisure therefore is low), agents will demand higher-quality leisure time, which means higher consumption during leisure periods. Also, note that current consumption depends negatively on labour supplied in $t+1$, since higher labour input tomorrow leads to a higher marginal utility of consumption, which would, on its own, cause households to sacrifice consumption today for consumption next period.

It is useful to recall that the ex ante real interest rate can be thought of as the price of a unit of consumption today in terms of a unit consumed next period. Higher interest rates induce individuals to postpone their consumption until later periods, to reap the benefits of a higher return on savings. Adding habit formation delays the peak response of consumption to movements in real interest rates as well as to the tax rate on consumption, while assuming an intertemporal elasticity of substitution of less than one reduces the sensitivity of consumption to interest rates and increases its sensitivity (in equal proportion) to movements in labour input.

Consumption by current-income consumers satisfies the no savings/dissavings condition that defines their behaviour:

$$P_{c,t}^{tot} C_t^c = \varkappa \left( (1 - \tau_{w,t}) W_t N_t + TF_t \right). \tag{2.52}$$

Regarding wage determination, we abstract for the moment from the particular utility function that we employ and focus instead on the intuition behind wage-setting behaviour in general. First, we note that, in demanding a higher wage, workers will enjoy higher labour income for the hours they work but will work less, since their relative wage will be higher. Of course, any reduction in hours worked will increase their leisure and hence their utility. These three effects can be seen directly through the Lagrangian that consumers maximize[22]:

$$\sum_{s=t}^{\infty} \boldsymbol{\beta}_{t,s} \left(\theta_w\right)^{s-t} \left\{ \underbrace{\mathbb{U}_s(\cdot)}_{a} - \Phi_s \left( [\circ]_s - \int_0^1 \underbrace{W_{ht}\Pi_{w,s,t}}_{b} \underbrace{\left( \frac{W_{ht}\Pi_{w,s,t}}{W_s} \right)^{-\epsilon_{w,s}}}_{c} H_s dh \right) \right\}, \tag{2.53}$$

where $a$ captures the direct effect of labour on utility. In other words, demanding a higher wage reduces the demand for one's labour input, thereby raising that individual's leisure, which enters utility positively. $b$ captures the direct effect of

[22] $[\circ]_s$ captures the other arguments in the consumer's budget constraint not related to the wage. $\Phi_s$ is the Lagrangian associated with the budget constraint.

higher wages on labour income, whereas $c$ captures the indirect effect of higher wages on the quantity of labour demanded. If the benefit of demanding a higher wage (captured by $a$ and $b$) outweighs the cost (captured by $c$), then the worker will indeed raise the wage they demand when renegotiating their nominal contract. Thus, at an aggregate level, agents will work an additional hour only if the extra consumption (captured by the price of consumption in terms of labour, $\frac{(1-\tau_{w,t})W_t}{P_{c,t}^{tot}}$), multiplied by the increase in utility per unit of extra consumption, exceeds the loss in utility from the extra hour worked. As a special case, with perfect competition in both the labour market and the domestic goods market, we obtain the efficient result that the marginal rate of substitution and the marginal product of labour are equated at the real wage.

The first-order condition for households that reset their nominal wage in period $t$ is given by

$$W_{h,t} = \mathbf{E}_t \frac{\frac{\mu-1}{\mu}\sum_{s=t}^{\infty}\beta_{t,s}(\theta_w)^{s-t}\mathbb{U}_s\left(E_s N_{hs}\right)^{\frac{1+\eta}{\eta}}\epsilon_{w,s}}{\sum_{s=t}^{\infty}\beta_{t,s}(\theta_w)^{s-t}\Phi_s(1-\varkappa)(1-\tau_{w,s})\Pi_{w,s,t}N_{hs}(\epsilon_{w,s}-1)}. \tag{2.54}$$

Linearizing the first-order condition for new wage contracts along with the equation describing the evolution of the aggregate wages,

$$W_t = \left(\theta_w\left(W_{t-1}\Pi_{w,t,t-1}\right)^{1-\epsilon_{w,t}} + (1-\theta_w)\left(W_{ht}^*\right)^{1-\epsilon_{w,t}}\right)^{\frac{1}{1-\epsilon_{w,t}}}, \tag{2.55}$$

yields the following equation for aggregate-wage inflation:

$$\begin{aligned}
\widehat{\pi}_{w,t} \quad = \quad & \frac{\gamma_w(1+r)}{1+r+g\gamma_w}\left(\widehat{\pi}_{w,t-1} - \Delta\mathbf{E}_t\overline{\pi}_{w,t}\right) + \frac{g}{1+r+g\gamma_w}\mathbf{E}_t\widehat{\pi}_{w,t+1} \\
& -\vartheta\left\{\widehat{w}_t - \left(\begin{array}{c}\frac{g\xi\beta(\mu-1)}{(g-\xi\beta)(g-\xi)\mu}\left(\Delta\widehat{c}_{t+1}^{l} - \xi\Delta\widehat{c}_t^{l} - \left(\frac{g-\xi}{g}\right)\overline{L}^{\frac{1+\eta}{\eta}}\Delta\widehat{L}_{t+1}\right) \\ +\frac{g}{(g-\xi\beta)(g-\xi)}\left(g\left(1+\xi^2\beta\right)\widehat{c}_t^{l} - \xi\widehat{c}_{t-1}^{l} - \xi\beta\widehat{c}_{t+1}^{l}\right) \\ +(1/\eta)\widehat{L}_t + \widehat{E}_t + \widehat{\tau}_{w,t}\end{array}\right)\right\} \\
& +u_{w,t},
\end{aligned} \tag{2.56}$$

with $\vartheta = \frac{\eta}{\eta+\epsilon_w}\frac{1-\theta_w}{\theta_w}\frac{(1+r-\theta_w g)}{1+r+g\gamma_w} > 0$. Similar to our pricing specifications, the dynamic wage equation contains a lag and a lead of wage inflation (relative to the perceived target for inflation adjusted for productivity growth), and an error-correction component. The presence of lagged wage inflation reflects the assumption of partial indexation of existing wage contracts to lagged wage inflation according to the following rule:

$$\Pi_{w,s,t} = \left(\frac{W_{s-1}}{W_{t-1}}\right)^{\gamma_w} \mathbf{E}_t \left(\prod_{j=t}^{s-1} (1+\overline{\pi}_j)\, g_j\right)^{1-\gamma_w} \quad s > t. \tag{2.57}$$

Whereas, in the pricing specifications, the long-run equilibrium condition was proportionality between price and marginal cost, here it is proportionality between the real wage and the marginal rate of substitution, adjusted for by effort and the labour tax.

To understand the mechanics of wage inflation, consider first the argument $\frac{1}{\eta}\widehat{L}_t$: when required to work more, consumers will demand higher wages. The extent will depend on the inverse of labour supply elasticity, $\eta$. The greater the loss in utility from a reduction in labour (the lower is $\eta$), the greater will be the wage increase required to compensate. Also note the presence of several consumption terms as well as $\Delta\widehat{L}_{t+1}$, which together capture the marginal contribution of a change in consumption to lifetime utility. Higher real consumption reduces the marginal value of consumption (the utility function is concave in consumption), thereby requiring a compensating increase in the wage to equilibrate it with the marginal utility of leisure. Consumers also consider their expected effort level ($\widehat{E}_t$) over the life of the contract. This stems directly from the fact that agents know that, while effort reduces their utility, there is no compensating increase in labour income (holding the wage fixed), because effort is not compensated; thus, agents will demand a wage premium to offset this effect (above and beyond what is already reflected in $\widehat{L}_t$). Finally, since households care about their wage only insofar as it allows them to purchase consumption goods, the appropriate wage measure is the nominal wage adjusted for by the direct tax rate on labour income, or the after-tax wage, $W_t(1-\tau_{w,t})$, divided by the after-tax price of consumption, $P_{ct}^{tot}$.[23] Thus, any increase in the labour income tax will eventually be matched one-for-one by an increase in the pre-tax real wage, all else being equal.

## 2.5 Foreign links

While Canadian primary-goods producers are assumed to operate in a perfectly competitive world market, we assume that Canadian exporters of finished products sell a good that is differentiated relative to its global competitors and therefore

[23] Recall that, in ToTEM, we assume a role for indirect taxes on consumption goods.

have some degree of market power. As a result, the rest-of-world demand curve for Canadian non-commodity exports is negatively sloped, and an increase in the supply of exports requires a depreciation of the value of the Canadian dollar. Demand for Canadian non-commodity exports is given by the following derived demand function:

$$\widehat{x}_{nc,t} = -\vartheta \widehat{p}^{*}_{m,t} + \widehat{y}^{*}_{t} + \vartheta_2 \widetilde{Y}^{*}_{t}, \tag{2.58}$$

where $\widehat{p}^{*}_{m,t}$ is the per cent deviation of the relative foreign-dollar price of Canada's non-commodity exports $\left(P^{*}_{m,t}/P^{*}_{t}\right)$, $\widehat{y}^{*}_{t}$ is the percent deviation foreign output (deflated by $A^{*}_{t}$)[24] relative to steady state, and $\vartheta$ is the elasticity of substitution between domestic non-commodity exports and foreign-produced goods. The last term is ad hoc and captures the impact on non-commodity exports of the rest-of-world output gap; the parameter $\vartheta_2$ is calibrated so that ToTEM can match the correlation between exports and rest-of-world GDP found in the historical data.

We also make an assumption for foreign import prices that is analogous to our assumption for domestic import prices, that foreign import prices are temporarily fixed in the foreign currency. The linearized equation for the foreign price of imports from Canada is

$$\begin{aligned}\widehat{\pi}^{m^*}_{t} \quad = \quad & \frac{\gamma_m}{1+\mathcal{B}\gamma_m}\left(\widehat{\pi}^{m^*}_{t-1} - \mathbf{E}_t \Delta \overline{\pi}_t\right) + \frac{\mathcal{B}}{1+\mathcal{B}\gamma_m}\mathbf{E}_t \widehat{\pi}^{m^*}_{t+1} \\ & + \frac{(1-\theta_m)(1-\mathcal{B}\theta_m)}{\theta_m(1+\mathcal{B}\gamma_m)}\left(\widehat{p}^{x}_{s} - \widehat{e}_t - \widehat{p}^{*}_{m,t}\right) + \varepsilon^{pm^*}_{t}.\end{aligned} \tag{2.59}$$

The consequence of this specification is that exchange rate movements feed into foreign import prices only gradually, thereby slowing the response of Canadian exports to movements in the exchange rate.

## 2.6 Monetary policy

For monetary policy, we consider two different calibrations of a simple, inflation-forecast rule. The first is calibrated so as to replicate the average behaviour of the monetary authority in Canada over the 1980–2004 sample, based on the estimates

[24] We also assume that the technology index for the foreign country, $A^{*}_{t}$, is cointegrated with $A_t$.

in Lam and Tkacz (2004) and Murchison (2001). This rule is used to compute the impulse responses presented in section 4.4. The second is a rule that prescribes an optimal path for interest rates conditional on a set of objectives for monetary policy, the structure of the model, and a particular form for the policy rule itself. This rule is used to complete the staff projection with ToTEM and is used for the shock presented in section 4.4.1, to illustrate differences in ToTEM's behaviour across the two rules.

### 2.6.1 *Historical rule*

Under this simple rule, monetary policy sets the nominal short-term interest rate, $R_t$, in response to deviations of current consumer price inflation, $\pi_t^c$, from $\pi_t^*$:

$$R_t = \Theta_R R_{t-1} + (1-\Theta_R)(\overline{r} + \overline{\pi}_t + \Theta_\pi(\pi_{t+2}^c - \pi_t^*)). \tag{2.60}$$

In addition, we assume, following Erceg and Levin (2003), that $\pi_t^*$ comprises stochastic components that vary through time:

$$\pi_t^* = \overline{\pi}_t + \pi_{qt} = \mathbf{H}\boldsymbol{\Upsilon}_t, \tag{2.61}$$

with $\mathbf{H} = \begin{bmatrix} 1 & 1 \end{bmatrix}$. $\overline{\pi}_t$ represents the permanent component or the inflation target and $\pi_{qt}$ represents transitory deviations from the rule, which we refer to as monetary policy shocks. Furthermore, agents in the economy are assumed to observe $\pi_t^*$ at time $t$ but not the decomposition implied by (2.61). Instead, agents solve a signal-extraction problem whereby they infer, at each point in time, the values of $\overline{\pi}_t$ and $\pi_{qt}$ based on their observation of the evolution of $\pi_t^*$. Specifically, the agents' estimate of this decomposition is updated based on their most recent forecast error of the target and the knowledge that these two unobservable components are driven by the process

$$\boldsymbol{\Upsilon}_{t+1} = \boldsymbol{\Psi}\boldsymbol{\Upsilon}_t + \boldsymbol{\varepsilon}_{t+1}: \quad \boldsymbol{\Upsilon}_{t+1} = \begin{bmatrix} \overline{\pi}_{t+1} \\ \pi_{q,t+1} \end{bmatrix}, \boldsymbol{\Psi} = \begin{bmatrix} \rho_p & 0 \\ 0 & \rho_q \end{bmatrix}, \ \boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_{p,t} \\ \varepsilon_{q,t} \end{bmatrix}, \tag{2.62}$$

but they do not directly observe $\boldsymbol{\Upsilon}_t$. Instead, given their knowledge of (2.62), and the assumption of orthogonality between $\varepsilon_{p,t}$ and $\varepsilon_{q,t}$, the agents update their beliefs according to the Kalman rule:

$$\mathbf{E}_t\boldsymbol{\Upsilon}_t = \boldsymbol{\Psi}\mathbf{E}_{t-1}\boldsymbol{\Upsilon}_{t-1} + \Sigma\left(\pi_t^* - \mathbf{H}\boldsymbol{\Psi}\mathbf{E}_{t-1}\boldsymbol{\Upsilon}_{t-1}\right). \tag{2.63}$$

Having computed their time $t$ expectation of the permanent and transitory components of the target, the agents use (2.62) to extrapolate the dynamic path of $\pi^*_{t+i}$ :

$$\mathbf{E}_t\pi^*_{t+i} = \mathbf{H}\mathbf{\Psi}^i\mathbf{E}_t\mathbf{\Upsilon}_t \qquad \forall i > 0. \tag{2.64}$$

For simplicity, we make the additional assumptions that $\rho_p = 1$ and $\rho_q = 0$.

### 2.6.2 *Optimized projection rule*

For the purpose of doing projections with ToTEM, we choose a monetary policy rule that minimizes an assumed loss function of the monetary authority (Cayen, Corbett, and Perrier 2006). We assume that the monetary authority has preferences over inflation stability with some concern for output stabilization relative to potential output. We also allow for the possibility that the authority cares about the volatility in the movements of its instrument. We formalize these preferences of the central bank with the following contemporaneous loss function:

$$L_t = (\pi^{tot}_t - \overline{\pi}_t)^2 + \lambda_y(\widehat{Y}_t)^2 + \lambda_{\Delta R}(\Delta R_t)^2, \tag{2.65}$$

with $\pi^{tot}_t$, $\widehat{Y}_t$, and $\Delta R_t$ being, respectively, the inflation rate in period $t$, the output gap (see chapter 4), and the change in the level of the policy instrument between period $t-1$ and period $t$. The parameters $\lambda_y$ and $\lambda_{\Delta R}$ are, respectively, the relative weights on output-gap and policy-instrument variability relative to inflation variability (around the target) in the preferences of the monetary authority. Some studies have justified setting $\lambda_{\Delta R} > 0$ on the basis that policy-makers may also care about the effect of instrument volatility on financial markets. Alternatively, we can appeal to the argument that $\lambda_{\Delta R} > 0$ reflects policy-makers' desire to avoid hitting the zero bound on nominal interest rates, a bound that does not exist in a linearized model such as ToTEM. In any case, we believe that a reasonable and sufficient justification is that, in reality, in any given period, the monetary authority (and other agents) is uncertain about the nature and the persistence of the shocks at play in the economy. Therefore, the monetary authority might want to avoid large policy reversals by smoothing its reaction to new developments.

The problem of monetary policy in period $t$ is to set its instrument in order to minimize the current and expected values for the period losses. More formally, we write the intertemporal loss function for the monetary authority as:

$$\mathcal{L}_t = (1-\beta)\mathbf{E}_t \sum_{i=0}^{\infty} \beta^i L_{t+i}, \tag{2.66}$$

with $\beta$ being the rate at which the central bank discounts future losses and $\mathbf{E}_t$, the conditional expectations operator, based on information available in period $t$. As is common in the literature and for operational purposes, we work with the unconditional version of (2.66), which is given by

$$\bar{\mathcal{L}} = \sigma_\pi^2 + \lambda_y \sigma_y^2 + \lambda_{\Delta R}\sigma_{\Delta R}^2 : \quad \lambda_y = 1.0, \lambda_{\Delta R} = 0.5, \tag{2.67}$$

where $\sigma_\pi^2$, $\sigma_y^2$, and $\sigma_{\Delta R}^2$ are the unconditional variances of the deviations of the inflation rate from the target, the output gap, and the change in the policy instrument, respectively.[25] The form of the reaction function that we consider is given as:

$$R_t = \Theta_R R_{t-1} + (1-\Theta_R)[\overline{r} + \overline{\pi}_t + \Theta_\pi(\mathbf{E}_t \pi_{t+h}^c - \overline{\pi}) + \Theta_y \widehat{Y}_t], \tag{2.68}$$

where $h$ is the monetary policy feedback horizon as described in Batini and Nelson (2001).

The objective of the exercise, then, is to choose $\{\Theta_R, \Theta_\pi, \Theta_y, h\}$ so as to minimize $\bar{\mathcal{L}}$, conditional on the structure of the model and the covariance matrix of structural shocks, which are calculated for the period 1992-2005 (see chapter 3 for a discussion of the parameter values chosen for the policy rule).

## 2.7 Government

The fiscal authority in ToTEM purchases the finished government product and distributes transfers to households, financing these expenditures through taxation and the issuance of debt. From the standpoint of private agents, no benefit is derived from government expenditures: households do not derive additional utility and government spending has no direct impact on the economy's productive capacity. While our set-up is quite simple, it does permit an analysis of the effects of several shocks, temporary or permanent, to the preferences of the fiscal authority.

[25] It can be shown that $\lim_{\beta\to 1}\mathcal{L}_t = \bar{\mathcal{L}}$ .

Fiscal policy in ToTEM is non-Ricardian for two reasons. First, as outlined previously, the existence of current-income consumers means that consumption responds directly to taxes and transfers. Second, we assume distortionary taxation in the form of a non-lump-sum tax on labour income. Any movement in this tax rate creates a substitution between leisure and consumption, and thus has affects on labour supply and consumption. These two features imply that fiscal policy can have a non-trivial impact on ToTEM's short-term dynamics.

The fiscal block contains behavioural equations for government spending, transfers, and direct and indirect income taxes, and a set of accounting identities that link these behavioural variables with the fiscal balance sheet. With the exception of the direct tax rate on labour income, fiscal behavioural variables are modelled as the sum of a discretionary and cyclical component. The discretionary component is set by the fiscal authority, and in theory changes only with new government policies,[26] while the cyclical component posits that fiscal policy responds to some cyclical variable, reflecting the stabilization role historically played by fiscal policy.

We begin by defining the budget constraint of the fiscal authority, who finances its deficit/surplus through the issuance of nominal bonds, $B_t$ :

$$B_{t+1} = (1 + R_t) \left( B_t + \left( P_t^g G_t + TF_t - TX_t \right) \right), \tag{2.69}$$

where $TF_t$ is the level of (nominal) transfers, and $TX_t$ is the level of revenues from all categories of taxes. The deficit/surplus includes interest payments on the public debt and the primary deficit, defined as spending minus taxes. The government obtains these tax revenues through a tax on labour income and on the consumption of goods and services:

$$TX_t = \tau_{w,t} W_t H_t + \tau_{c,t} \left( \frac{P_{c,t}^{tot} C_t^{tot}}{1 + \tau_{c,t}} \right), \tag{2.70}$$

where $\tau_{w,t}$ is the (direct) tax rate on labour income and $\tau_c$ is the indirect tax rate on consumer expenditures, whose behaviour is governed by a simple autoregressive stochastic process of the form

$$\tau_{c,t} = \Phi_c \tau_{c,t-1} + \varepsilon_t^{\tau_c} \qquad \varepsilon_t^{\tau_c} \sim iid(0, \sigma_{\tau_c}^2). \tag{2.71}$$

The government's main objective is to achieve its desired debt-to-GDP ratio,

[26] In practice, the discretionary component is the difference between the model's prediction and the historical data.

$\left(\overline{\frac{B_t}{P_t Y_t}}\right)$, over the medium term. To this end, a fiscal policy reaction function links $\tau_{w,t}$ to the deviation of the actual debt-to-GDP ratio from its desired level:

$$\tau_{w,t}= \Phi_{w,1}\tau_{w,t-1}+(1-\Phi_{w,1})\left(\tau_w + \Phi_{w,2}\left(\frac{B_{t+1}}{P_{t+1}Y_{t+1}} - \left(\overline{\frac{B_t}{P_t Y_t}}\right)\right)\right)+u_t^{\tau_w}, \quad (2.72)$$

where $\tau_w$ is the steady-state value for the tax rate on labour income, and $u_t^{\tau_w}$ is a temporary tax shock

$$u_t^{\tau_w}= \rho_{\epsilon_{\tau_w}} u_{t-1}^{\tau_w}+\varepsilon_t^{\tau_w} \qquad \varepsilon_t^{\tau_w}\sim iid(0,\sigma^2_{\tau_w}).$$

The fiscal policy instrument, $\tau_{w,t}$, is distortionary, in that it affects the marginal return to supplying labour, and hence the amount of labour that a household will agree to supply at a given pre-tax wage rate. Furthermore, any movement in a non-distortionary fiscal variable, such as the share of transfers in GDP, will also generate distortionary effects, since its impact on the level of debt ultimately must be offset through a change in the direct tax rate.

The fiscal block incorporates two types of government spending: purchases of goods and services, and transfers to households. We have chosen to model these two components of spending separately, because of the distinct roles they play in the model. Specifically, the assumption that some consumers are credit constrained implies that transfers, which enter the household's budget constraint, can have an important influence on consumption spending. On the other hand, as previously mentioned, government spending on its own has no value for economic agents.

For government spending on goods and services, we specify the following equation:

$$\left(\frac{P_t^g G_t}{P_t Y_t}\right) = \Psi_g\left(\frac{P_{t-1}^g G_{t-1}}{P_{t-1}Y_{t-1}}\right)+(1-\Psi_g)\left(\overline{\frac{P_g G}{PY}}\right)+u_t^g, \quad (2.73)$$

where $u_t^g$ follows the autoregressive process:

$$u_t^g = \rho^g u_{t-1}^g + \varepsilon_t^g \qquad \varepsilon_t^g \sim iid(0,\sigma_g^2). \quad (2.74)$$

For the share of transfers to nominal GDP, we specify the identical form:

$$\left(\frac{TF_t}{P_t Y_t}\right) = \Psi_{tf}\left(\frac{TF_{t-1}}{P_{t-1}Y_{t-1}}\right)+(1-\Psi_{tf})\left(\overline{\frac{TF}{PY}}\right)+\varepsilon_t^{tf} \qquad \sim iid(0,\sigma_{tf}^2). \quad (2.75)$$

## 2.8 Additional market-clearing conditions

ToTEM is closed with the following market-clearing conditions that define GDP ($Y_t$) and the GDP deflator ($P_t$):

$$P_t Y_t = P_{c,t}^{tot} C_t^{tot} + P_t^I I_t + P_t^g G_t + P_t^x X_{nc,t} + e_t \left( P_{com,t}^* X_{com,t} - P_t^* M_t \right) \quad (2.76)$$
$$Y_t = C_t^{tot} + I_t + G_t + X_{nc,t} + X_{com,t} - M_t. \quad (2.77)$$

In terms of the inputs to production, the following conditions ensure that the sums of the sectoral inputs equal the total allocations:

$$\begin{aligned} L_t &= L_t^c + L_t^I + L_t^g + L_t^{nx} + L_t^{com}, \\ H_t &= H_t^c + H_t^I + H_t^g + H_t^{nx} + H_t^{com}, \\ K_t &= K_t^c + K_t^I + K_t^g + K_t^{nx} + K_t^{com}, \\ I_t &= I_t^c + I_t^I + I_t^g + I_t^{nx} + I_t^{com}, \\ M_t &= M_t^c + M_t^I + M_t^g + M_t^{nx}. \end{aligned}$$

Aggregate effort, $E_t$, is equal to a weighted average of the effort levels in each sector (using the market-clearing condition for $L_t$ along with the assumption that $L_t = H_t E_t$).

Finally, combining the budget constraints for the two types of households with that of the government, and using the definitions of dividends, we obtain the following equation describing the evolution of the net foreign-asset position denominated in Canadian dollars:

$$e_{t+1} B_{t+1}^* = \frac{e_{t+1}}{e_t} \left(1 + R_t^*\right) \left(1 + \kappa_t\right) \left\{ e_t B_{h,t}^* + P_t^x X_{nc,t} + e_t P_{com,t}^* X_{com,t} - e_t P_t^* M_t \right\}. \quad (2.78)$$

# Chapter 3

# Model Calibration

## 3.1 Why calibration?

Given the recent literature on the formal estimation of medium-scale DSGE models, it would seem at first glance that the estimation of ToTEM would be relatively straightforward. But in attempting to estimate ToTEM, the staff have encountered several problems, some of which are conceptual in nature while others are more practical. In previous work using an early prototype of ToTEM, Murchison, Rennison, and Zhu (2004) follow CEE in choosing parameters to minimize the difference between the impulse responses from a vector autoregression (VAR) and those generated by the model. This approach holds a certain intuitive appeal, since model behaviour at the Bank is often characterized in terms of responses to shocks around some control simulation. In addition, the flexibility to focus on some impulse responses while omitting others is generally viewed as a benefit over methods such as maximum likelihood. However, this approach is not without its problems. Since our work in 2004, ToTEM has grown substantially in size to incorporate the level of detail required of a projection and policy-analysis model. Our experience suggests that, as the number of parameters increases beyond a relatively small set, one cannot say with much confidence that the optimization routines are converging on a global minimum. Practical experience suggests that the estimation results are very sensitive to the set of starting values chosen.

In addition to the practical problems with the impulse-response-matching approach, there is also much disagreement regarding how best to identify the structural shocks in the VAR. Identifying a VAR of realistic dimensions requires the imposition of a large number of untestable restrictions. Typically, these restrictions take the form of zeros in the time$-t$ impulse response matrix. However, these restrictions are generally inconsistent with the structure of our model, at least when it is solved under the assumption of time$-t$ rational expectations, which is our preferred specification. The lagging of expectations in the structural model to make it consistent with a set of untestable restrictions imposed on the VAR is also questionable, on the grounds that it does not address the fundamental issue that a large set of arbitrary restrictions is being imposed on the model, in this case both the VAR and ToTEM.[27]

A more general difficulty faced by the staff in estimating ToTEM stems from the multiple criteria that a policy-analysis and projection model at the Bank is expected to satisfy. ToTEM's role as a projection model implies that a large weight will be placed on its forecasting ability. But when examining individual shocks, the Bank staff's strong preference for certain features effectively ties the hands of the modeller and makes formal estimation extremely challenging. Often, reduced-form relationships such as the peak impact of the exchange rate on output and inflation that have been estimated by staff in the past are viewed as crucial features for the model to match. In many cases, it is extremely difficult to formally write down these criteria in an objective function that can be explicitly maximized as one could in, for example, a strict moment-matching exercise. Clearly, much work in the field has gone into the development of routines, such as Dynare, that allow the incorporation of priors on the set of structural parameters. However, our experience indicates that the priors held by Bank staff are more closely related to the behaviour of the model, than to the structural parameters themselves.

For the time being, we have opted to pursue an informal parameterization strategy with no explicit objective function. Parameters have been chosen with a goal of matching a set of univariate autocorrelations, bivariate cross-correlations, and variances estimated using detrended historical data from 1980-2004. We have also used information gleaned from empirical studies such as Amano and van Norden (1995) (elasticity of the real exchange rate with respect to the terms of trade), and from estimated reduced-form models such as NAOMI (see Murchison 2001), Duguay (1994), and the VAR estimated in Murchison, Rennison, and Zhu (2004).

[27] An interesting avenue for future research would be to impose a set of just-identifying restrictions on the VAR, derived from the structural model.

### 3.1.1 *Steady-state parameters*

We partition the parameters into two groups: "steady state" and "dynamic." In the former group are parameters whose values are chosen primarily such that the model will have certain desired steady-state properties. For example, our chosen value for a household's discount rate $\beta$ respects the model's steady-state stability condition,

$$\beta = (1 + r)^{-1} g^{\frac{1}{\mu}},$$

conditional on choices for the economy's steady-state per-capita quarterly growth rate, $g = 1.005$, and the real quarterly interest rate, $r = 0.008$. Other examples include the distribution parameters ($\alpha$) that appear in each production function in the model. These are chosen so that the steady-state model replicates the historical averages for relative shares (or great ratios) found in the data. These ratios are then used to linearize the dynamic model. The parameter $\rho$ is chosen such that capacity utilization is equal to one in steady state.

### 3.1.2 *Dynamic parameters*

We turn next to ToTEM's dynamic parameters. We first focus on dynamic parameters associated with nominal rigidities in the model. The parameter $\gamma_c$, which determines the extent to which non-reoptimizing firms in the consumption-goods producing sector can index to lagged inflation, is set to 0.18 ($\gamma_c \in [0, 1]$). This value is somewhat lower than the 0.4 reported in Amano and Murchison (2005). It was chosen primarily to help the model replicate the moderate level of inflation persistence estimated in the data from 1980-2004. Inflation in ToTEM inherits a substantial amount of persistence from real marginal cost, and therefore a high degree of indexation is not required. This setting suggests that the weight on the forward-looking component is quantitatively more important than is the weight on lagged inflation (0.8 versus 0.2). The indexation parameters in the investment, government, and export sectors are also set to 0.18.

The parameter $\theta_c$, which determines the proportion of firms that are not chosen to reoptimize every period, is 0.7, implying that domestic price contracts are reoptimized, on average, once every three quarters. This value is also used for the other finished-product sectors.

For import-price inflation, the dynamic-indexation parameter, $\gamma_m$, and the contract length parameter, $\theta^m$, are both 0.9, implying that import-price contracts are reoptimized, on average, once every 10 quarters. The long length of time between optimizations and the high degree of indexation reflect the relatively low and gradual short-run exchange rate pass-through seen in the data over our sample period. The same parameterization governs the price of Canadian exports in U.S.-dollar terms.

For wages, the degree of indexation, $\gamma_w$, is 0.18 and the probability of not being picked to reset, $\theta_w$, is 0.85, meaning that wages are reoptimized about every 6.5 quarters, on average. Thus, the main source of nominal rigidity in ToTEM is sticky wages, consistent with the survey evidence presented in Amirault, Kwan, and Wilkinson (2006) for prices, and in Longworth (2002) for wages.

In ToTEM, household consumption depends positively on lagged consumption according to the habit-persistence parameter, $\xi$, which we set to 0.65, as in CEE. We set the intertemporal elasticity of substitution, $\mu$, to 0.9, within the range of 0.5 to 1 made in much of the literature on real business cycles, while the wage elasticity of labour supply, $\eta$, is set equal to 0.6. We set the share of credit-constrained consumers equal to 20 per cent, which allows the model to generate a reasonable impact of government debt shocks on output.

We turn next to our parameterizations for the model's various adjustment costs. We set $\chi_k$ ($\chi_u$), which is the cost, in terms of the output of the final good in each sector, associated with changing the level of capital (investment), to 10 (30). The costs of varying capital utilization are governed by the parameter $\chi_u$, which we set at 5.0; this setting implies that adjusting capital utilization is relatively expensive or, alternatively, that the marginal benefits of increasing utilization are quite small. The labour adjustment-cost term, $\chi_L$, is set to 1.5, which yields a modestly procyclical (counter-cyclical) profile for labour productivity (labour share of income) for most demand-style shocks.

The elasticities of substitution between the various inputs to production ($\sigma, \varphi, \varrho$) are currently set to 0.5, following the work of Perrier (2005) (this also the value used for Canada in Gagnon and Khan 2005). While we have the flexibility to set different values for these three parameters, the informal calibration method that we have employed thus far does not really permit us to identify separate roles for each. We expect this will change once we formally estimate the model.

Given the calibrations of the adjustment-cost parameters for capital, investment, and capital utilization, as well as the elasticities of substitution in production, the hazard rate $\theta_c$, and the elasticity of substitution between finished goods in the consumption sector, $\zeta$ is equal to 0.25, meaning that aggregate CPIX inflation is four-times less sensitive to real marginal cost than if we were to assume a homogeneous capital market. Taken together, $\theta_c$ and $\zeta$ suggest a short-run elasticity of inflation with respect to a real marginal cost of 0.03, which lies between the value of 0.04 reported for the United States in ACEL (see also Eichenbaum and Fisher 2004), and the valid of 0.02 reported for Canada in Gagnon and Khan (2005).[28] Much of the persistence of inflation in ToTEM is generated by the assumption of firm-specific capital that is costly to adjust, in conjunction with a small amount of nominal rigidity in the goods sectors, rather than by nominal rigidities alone.

In the hybrid-UIP condition, we set the weight on the lagged nominal exchange rate, $\varpi$, to 0.48. This value was chosen with the goal of both replicating the unconditional moments of the historical data and of generating sensible impulse responses. The chosen value reduces the volatility of the exchange rate, delays its peak response typically by about two quarters, and increases the persistence of the real exchange rate.

For the fiscal policy rule that determines the direct tax rate on labour income, equation (2.72), we set the weight on the lagged direct tax rate, $\Phi_{w,1}$, to 0.7, and the weight on the deviation of the debt-to-GDP ratio from its target, $\Phi_{w,2}$, to 1. For the behavioural equation for government spending, we set the weight on the lagged dependent variable, $\Psi_g$ in equation (2.73), to 0.94, which implies a weight on the optimal steady-state ratio of 0.06. This setting implies that nominal government expenditures on goods and services adjust very gradually to movements in nominal GDP, which allows the model to capture the high degree of autocorrelation found in the business-cycle dynamics of government spending.

As for monetary policy in ToTEM, while it is clear that no single rule could adequately characterize monetary policy between 1980 and 2004, we choose a parameterization that we feel best reflects the average behaviour of monetary policy over history. Using the parameter estimates reported in Cayen, Corbett, and Perrier (2006), we set the smoothing parameter, $\Gamma_r$, to 0.8 and the weight on inflation deviations from target, $\Gamma_\pi$, to 2.5. Finally, the parameters for the optimized inflation-forecast monetary policy rule are given as $\{\Theta_R = 0.95, \Theta_\pi = 20, \Theta_y = 0.35, h = 2\}$, with the short-run response coefficients given as $(1 - \Theta_R)\,\Theta_\pi = 1$ and $(1 - \Theta_R)\,\Theta_y = 0.02$. It is interesting to note the high value for the smoothing parameter $\Theta_R = 0.95$, which is optimized over the range $\Theta_R \in [0, 1)$. This reflects a combination of the crucial role played by the expectations of future outcomes in the model and an

[28] These authors also assume CES production with an elasticity of substitution of 0.5 when computing real marginal cost.

assumed desire on the part of the monetary authority in the model to reduce unnecessary instrument volatility. Essentially, because the model is so forward looking, monetary policy can achieve nearly the same output/inflation outcome in response to a shock by moving interest rates by a great deal for a short period of time or by a lesser amount for a long period of time. Given the presence of $\Delta R_t$ in the loss function, the latter is the preferred outcome.

$\Sigma$ is chosen to replicate a sacrifice ratio of around 2 (which corresponds to the average estimate in the literature for Canada), as well as to replicate the speed of adjustment of private sector inflation expectations during periods of disinflation in Canada in the early 1980s and 1990s.

**Table 2: ToTEM's Key Behavioural Parameters**

| Coefficient | Value | Coefficient | Value |
|---|---|---|---|
| *Manufacturing sector (section 2.1)* | | *Households (section 2.4)* | |
| $\sigma, \varrho, \varphi$ | $0.5$ | $\theta_w$ | $0.85$ |
| $\chi_k, \chi_I, \chi_u, \chi_h$ | $10, 30, 5.0, 1.5$ | $\gamma_w$ | $0.18$ |
| $g, r$ | $0.005, 0.008$ | $\varkappa$ | $0.2$ |
| $\epsilon, \epsilon_w$ | $21, 21$ | $\rho_\beta, \rho_\kappa, \varpi$ | $0.71, 0.66, 0.48$ |
| $\zeta$ | $0.2$ | $\varsigma$ | $0.009$ |
| $\omega$ | $0.016$ | $\mu, \eta, \xi$ | $0.9, 0.6, 0.65$ |
| $\theta_c, \theta_I, \theta_g, \theta_{X_{nc}}$ | $0.7$ | *Foreign sector (section 2.5)* | |
| $\gamma_c, \gamma_I, \gamma_g, \gamma_{X_{nc}}$ | $0.18$ | | |
| $\rho_\epsilon^c, \rho_\epsilon^I, \rho_\epsilon^g, \rho_\epsilon^{X_{nv}}$ | $0.73, 0.23, 0.5, -.13$ | $\delta$ | $1.2$ |
| $\psi$ | $0.5$ | $\vartheta$ | $1.5$ |
| *Import/Commodity sector (section 2.2)* | | *Monetary policy (section 2.6)* | |
| $\theta_m, \theta_{com}$ | $0.92, 0.79$ | $\Sigma$ | $0.22$ |
| $\gamma_m, \gamma_{com}$ | $0.9, 0.2$ | $\Theta_R, \Theta_\pi$ | $0.8, 2.5$ |
| $\chi_h$ | $5.0$ | *Government (section 2.7)* | |
| $\upsilon$ | $0.2$ | $\Phi_c, \Phi_{w,1}, \Phi_{w,2}, \rho_{\epsilon_{\tau_w}}$ | $0.95, 0.7, 1.0, 0.8$ |
| | | $\Psi_{tf}, \Psi_g$ | $0.95, 0.94$ |

# Chapter 4

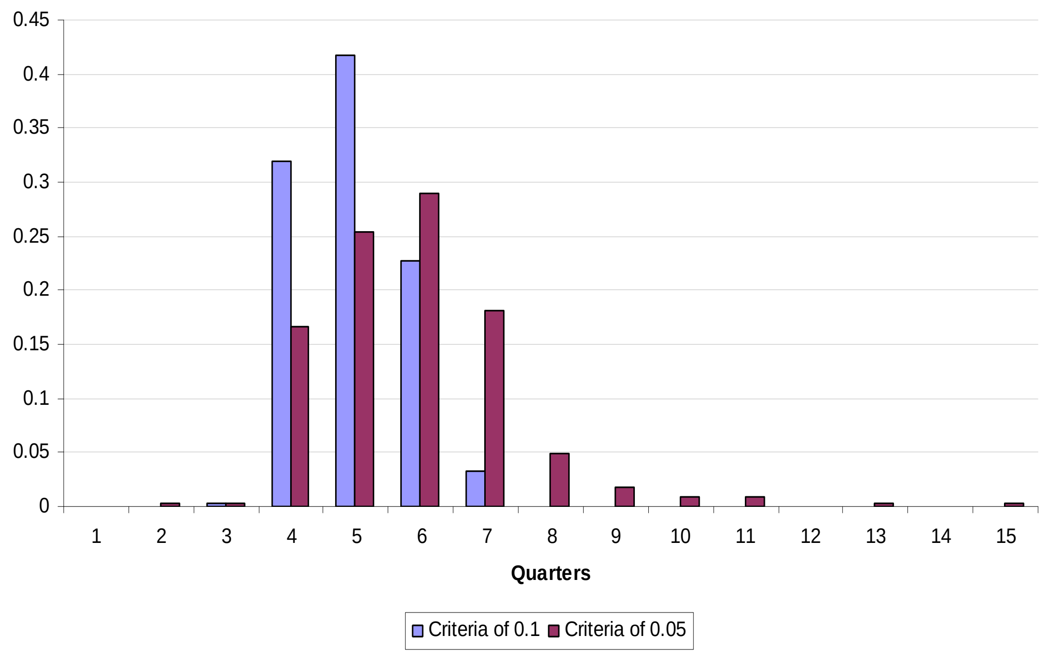
# Model Properties

## 4.1 The target and feedback horizons

The time-series properties of inflation have changed since the beginning of the 1990s (see Longworth 2002 for an extensive review). While the exact date on which these changes occurred is subject to some uncertainty, it is nevertheless the case that both the volatility and the persistence of inflation have declined markedly in the 1990s and 2000s, relative to previous decades. In addition, the slope of the empirical Phillips curve has decreased, as has the extent to which exchange rate movements get passed through to the CPI. In other words, inflation is now less sensitive to excess demand and supply pressures, as well as to movements in relative prices such as the exchange rate.

These changes in the properties of inflation are reflected in the behaviour of ToTEM, and have direct implications for both the inflation-target horizon and the optimal feedback horizon (Batini and Nelson 2001). The inflation-target horizon is defined as the average length of time it takes for inflation to return to the target. While this value depends importantly on the preferences of the central bank, it also depends on the parameterization of the rest of the model. For instance, a low degree of indexation in prices and wages ($\gamma_c$ and $\gamma_w$ close to zero) will, all else being equal, reduce the persistence of inflation and, consequently, reduce the target horizon. A high degree of nominal rigidity in the model, by contrast, will tend to increase the persistence of inflation. Using ToTEM's current calibration, it takes seven quarters, on average, for inflation to return to within $\pm 0.1$ percentage points of the target following a shock in ToTEM, compared with about 10 quarters in QPM.[28] Moreover, if the sample from which the shocks are drawn is restricted to the less-volatile 1992-2005 period (Figure 3), the average time declines to five quarters for the ±0.1 percentage-points band, and six quarters for the ±0.05 percentage-points band.

**Figure 3: Optimal Inflation-Target Horizon with Optimized Policy Rule**




The second implication of lower inflation persistence is a shorter monetary policy feedback horizon, which corresponds to $h$ in the monetary policy reaction function:

$$R_t = \Theta_R R_{t-1} + (1-\Theta_R)[\overline{r} + \overline{\pi}_t + \Theta_\pi(\mathbf{E}_t \pi^c_{t+h} - \overline{\pi}) + \Theta_y \widehat{Y}_t].$$

In ToTEM, monetary policy need not look as far into the future when setting policy, since, all else being equal, the maximum impact on inflation of monetary policy actions arrives sooner. This implication is reflected in the calibration of

[28] See Cayen, Corbett, and Perrier (2006) for more details.

ToTEM's optimized monetary policy rule; when specified in terms of year-over-year (quarterly) inflation, the relevant policy feedback horizon is four (two) quarters, which is, again, about half of the six to eight quarters assumed in QPM for year-over-year inflation.

## 4.2 Exchange rate pass-through

In Canada, roughly one-quarter of the CPI basket comprises imports and, as such, the issue of pass-through of the exchange rate to the CPI is quite central. Historically, the issue of pass-through has centred on the question of whether movements in the exchange rate affect the price level or the inflation rate in the long run (Duguay 1994). However, more recent data support the hypothesis that exchange-rate movements affect the inflation rate in the short run only, so the price level alone is affected in the medium term.

In a model such as ToTEM, the extent to which the exchange rate gets passed through to prices will depend on many factors, not the least of which is the monetary policy reaction function. Through its effect on inflation expectations, policy can have a large influence on the short-run inflation response to any relative-price change. All else being equal, an aggressive response by an inflation-targeting central bank will tend to reduce measured pass-through when the policy response is known and well understood by private agents. In the context of the Phillips curve (equation (2.29) from section 2.1), monetary policy will act to reduce the expected persistence of the real marginal-cost response stemming from an exchange rate movement. In other words, for a given exchange rate movement at time $t$, the absolute value of the sum

$$\sum_{i=0}^{\infty} \left(\mathcal{B}^{i}\right) \widehat{\lambda}_{t+i}^{c},$$

will be lower when the policy response is more aggressive.

Some of ToTEM's structural parameters will also play an important role in the determination of short-run pass-through. For instance, recall that the coefficient that equates current core CPI inflation to current and future marginal costs (see equation (2.29)) is given by

$$\frac{\zeta \left(1-\theta_{c}\right)\left(1-\theta_{c} \mathcal{B}\right)}{\theta_{c}}.$$

Thus, for a *given* expected sequence of real marginal costs, the short-run inflation response is increasing in $\zeta$, which is governed by our assumptions about the structure of the market for capital, and decreasing the average contract length (determined by $\theta_c$). Finally, the weights in the marginal-cost expression, which reflect the importance of the exchange rate in determining factor input prices, will play a role in the response of real marginal cost. Recall that real marginal cost in the core consumption sector is given by the expression:

$$\widehat{\lambda}_t^c = \omega_1 \widehat{w}_t + \omega_2 \widehat{p}_{com,t} + \omega_3 \widehat{p}_{m,t} + \omega_4 \widehat{p}_{inv,t} + \omega_5 \widehat{u}_t,$$

where the weights ($\omega_i$) mainly reflect the relative importance of each factor in the production of consumption goods. Exchange rate movements will have a direct effect on the relative price of imports, commodities, and investment.[29] In addition, the real wage will be affected indirectly, since higher consumer prices eventually lead to demand for higher wages as workers attempt to maintain a particular real wage. The magnitudes of each of these influences will be determined by the weights $\{\omega_1, \omega_2, \omega_3, \omega_4\}$.

How large, then, is exchange rate pass-through in ToTEM? To answer this question, one must first decide on how pass-through should be measured. In empirical studies, it is customary to first estimate a reduced-form Phillips curve on historical inflation data that includes a role for the exchange rate or import prices. Pass-through at various horizons can be computed directly from the estimated parameters of the equation. For our purposes, it is useful to have a measure that can be calculated in an equivalent manner in both ToTEM and QPM.[30] Thus, we define pass-through simply as the percentage change in the core consumer price level at a particular time horizon stemming from an initial 1 per cent, exchange rate movement that arrives at time $t$:

$$\frac{\ln(P_{t+k}/P_{t-1})}{\ln(e_t/e_{t-1})},$$

where $k$ is the time horizon of interest. Note here that we divide by $\ln(e_t/e_{t-1})$ and not $\ln(e_{t+k}/e_{t-1})$. In both models, $\frac{\ln(P_{t+k}/P_{t-1})}{\ln(e_{t+k}/e_{t-1})} \to 1$ as $k$ grows large, and is

[29] Recall that investment goods have a high import concentration, and that therefore their price is heavily influenced by the price of imports.

[30] We could generate stochastic data using both models and then estimate reduced-form Phillips curves on the artificial data. However, the results would depend strongly on the covariance matrix of shocks assumed for each model. While it is straightforward to compute this covariance matrix for ToTEM using historical data, it is very difficult and time consuming in QPM.

therefore less useful as a pass-through measure.

Using this definition and $k = 3$ (the end of year one), both QPM and ToTEM predict pass-through of about 0.05 per cent. After two years, however, QPM predicts that this number rises to 0.18 per cent, about double that of ToTEM, and the difference continues to grow with the time horizon.[31]

## 4.3 The transmission of monetary policy

Monetary policy in ToTEM is able to affect real activity and relative prices because prices and wages are not fully flexible in the short run. As discussed in section 1.5.2, monetary policy influences directly the nominal short-term interest rate, $R_t$, whereas consumption and investment are driven by the ex ante real interest rate, which can be expressed as:

$$\widehat{r}_t \equiv \widehat{R}_t - \mathbf{E}_t \widehat{\pi}_{t+1}.$$

The inflation rate in period $t + 1$ will vary across sectors, since the relevant price index for determining $\widehat{\pi}_{t+1}$ is the price of that sector's output. In addition, the relevant price index for consumption is the total CPI ($\widehat{\pi}_{t+1}^{tot}$). With nominal rigidity, movements in the nominal interest rate cause movements in the real rate precisely because $\mathbf{E}_t \widehat{\pi}_{t+1}$ cannot adjust sufficiently to keep the real rate constant. Therefore, monetary policy, as measured through the 90-day real interest rate, exerts a direct impact on consumption and investment, and an indirect impact on net exports and government expenditures.

In the case of consumption, the ex ante real interest rate is the real price of consumption today, measured in units of consumption in the next quarter. Therefore, policy can introduce an intertemporal substitution effect in consumption. With habit formation, however, consumers are assumed to grow accustomed to a certain consumption level and therefore do not like large variations in their consumption spending. Consequently, the impact of a change in interest rates on consumption builds gradually through time. These changes also work their way through to the lifetime-income consumer's budget constraint, and create a small income effect (of the opposite sign). For current-income consumers, the effect of monetary policy

[31] A qualitatively similar result is obtained for differences in the influence of excess demand (or supply) on inflation across the two models. In general, a shock to domestic demand causes inflation to rise by less, and the peak response occurs sooner and diminishes faster in ToTEM relative to QPM.

occurs indirectly, through its influence on fiscal policy. Expansionary monetary policy, for instance, will lead to an increase in government revenues and a reduction in the government debt-to-GDP ratio at the existing tax rate. In an effort to restore the debt ratio at the government's desired level, the direct tax rate on labour income will decline, causing current-income consumers to raise their consumption.

On the investment side, the real interest rate is the rate at which firms discount future real profits (one quarter in the future) accruing from a change to their current capital stock. Higher real interest rates therefore reduce the net present value of future profits stemming from investment spending today (reduce $q_t$), and therefore reduce investment demand. In addition, monetary policy affects investment indirectly both through its influence on the marginal product of capital (the policy affects overall demand) and through its influence on the relative price of investment (mainly through the exchange rate). Because of adjustment costs, however, the initial response to a change in interest rates is small and builds gradually through time, creating a "hump-shaped" response that peaks after about five quarters for a typical interest rate change.

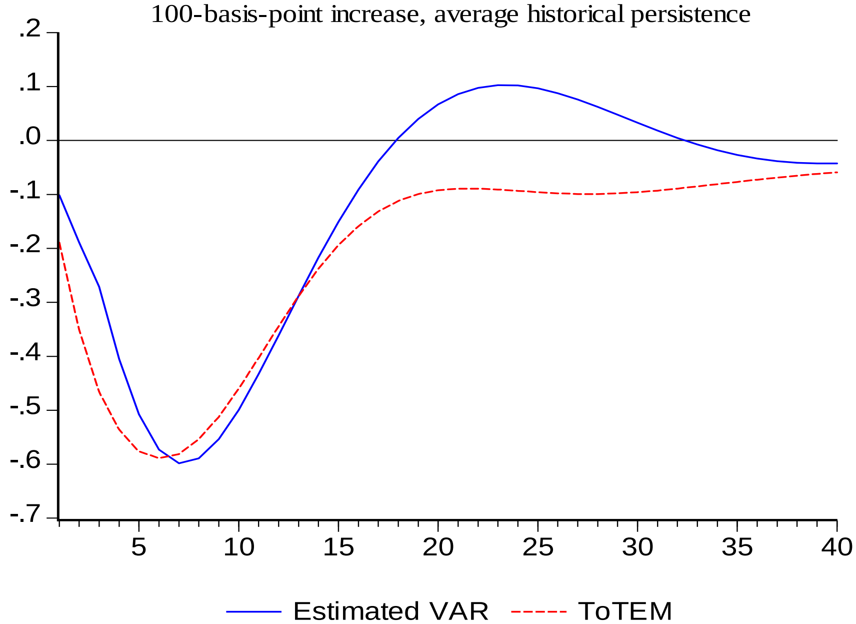
On the trade side, manufactured and commodity exports are influenced through monetary policy's impact on the exchange rate and, ultimately, the price paid by foreigners. Imports, in turn, are affected via this exchange rate effect (substitution effect), as well as by changes in demand for finished products such as consumption and investment goods. Finally, government expenditures rise in response to the increase in GDP, to maintain the desired share in government spending.

Figure 4 shows the impact on real GDP of a 100-basis-point increase in nominal short-term interest rates that is run off over the course of about two years, using ToTEM's current calibration. Also plotted is the response of real GDP from an identified VAR model for the same interest rate shock, taken from Rennison, Murchison, and Zhu (2004), which is estimated from 1980 to 2003. Overall, ToTEM does a good job of replicating the timing and magnitude of the GDP response. While ToTEM does predict that GDP should decline somewhat faster than the VAR in the first year, the differences are not large. Furthermore, ToTEM predicts that the peak decline in output should arrive in quarter five, whereas the VAR predicts a trough in quarter six. Again, this difference is small and would certainly fall within the 5 and 95 percentile ranges of the VAR.

Overall, the response of GDP in ToTEM is quite close to that of the VAR for the first four years, after which ToTEM predicts weaker GDP. This stems from our assumption of habit formation in consumption and adjustment costs on changes in investment. On the investment side, for instance, firms are reluctant to quickly increase their investment back to pre-shock levels, because big swings in investment reduce their overall productivity.

**Figure 4: Output Response to Interest Rates**

100-basis-point increase, average historical persistence




## 4.4 Impulse responses

We turn next to a discussion of ToTEM's impulse responses, which are computed as the per cent deviation of each variable (percentage or basis-point deviation for rates of growth such as inflation and interest rates) from its control path at a quarterly frequency. Since simulations are carried out using a log-linearized version of ToTEM, the starting-point conditions are of no consequence to the shock-control results.[32] All simulations, except for the "optimized-rule" scenario in the demand shock, are performed using the historically estimated monetary policy rule with weights given by $\{\Theta_R = 0.8, \Theta_\pi = 2.5, \Theta_y = 0.0, h = 2\}$ (details are given in the previous section).

[32] With the exception of the second disinflation experiment.

In addition to the response of several of ToTEM's key behavioural variables, we include for some simulations an output-gap measure that is conceptually similar to that used in QPM. While this variable plays no behavioural role in ToTEM,[33] it is nevertheless a useful summary variable for the state of excess demand or supply. As shown in Appendix B, the output gap in ToTEM can be represented as a weighted average of the labour and capital-market gaps in each sector of the economy. More specifically, the output gap is defined to be:

$$\widehat{Y}_t \equiv \sum_{j=1}^{5} \zeta_j \widehat{L}_{j,t} + \sum_{j=1}^{5} \rho_j \widehat{u}_{j,t},$$

where $\widehat{L}_{j,t}$ and $\widehat{u}_{j,t}$ are, respectively, the labour input and capital utilization gaps in each of the finished-product sectors (indexed by $j$), which include the consumption, investment, government, and commodity and non-commodity export sectors.

### 4.4.1 *A demand shock*

In this section, we discuss the results of an exogenously driven increase in private consumption, driven by a temporary reduction in the households' discount rate, $\beta_t$ (see section 3.1.1), which causes consumption to increase by about 1.25 per cent at the end of the first year of the shock. In addition to the results obtained using an historically estimated monetary policy rule, we present the interest rate and inflation responses using the optimized rule from section 2.6.2.

We begin with the composition of domestic demand (Figure 5). A common problem associated with this class of models is they predict that consumption and investment should move in opposite directions following a shock to demand, which is counterfactual. Typically, following an increase in consumption, real interest rates rise sufficiently to more than offset the impact of higher expected demand on the marginal product of capital ($q_t$ declines). However, given ToTEM's multi-goods set-up, the relative price of investment, which is heavily influenced by a real appreciation of the currency given investment's high import concentration, declines by even more, generating a small increase in demand for capital goods. A peak increase of just over 0.1 per cent occurs in the middle of the second year of the simulation.

The initial impact of the shock is an increase in real GDP of about 0.45 per cent

[33] Apart from appearing in the optimized reaction function with a very small weight.

by the end of the first year, after which output gradually returns to control. The combination of stronger demand for consumption goods, which requires imports as factor inputs, and a 0.6 per cent real appreciation, generates a 0.7 per cent increase in import demand, while exports fall by about 0.5 per cent. Thus, while GDP increases following a shock to domestic demand, the trade balance worsens, suggesting that some of the extra consumption is borrowed from abroad.

In ToTEM, an unexpected increase in demand for consumption goods is met by firms in the short run through an increase in the use of variable inputs, which consist of labour, capital services, commodities, and imports. While the firm chooses the input combination to minimize their costs, there exists no combination that will allow them to increase their production without increasing marginal cost, even if the price of factor inputs remains unchanged. For instance, if the firm attempts to increase all four factors in equal proportion, which would seem to be a reasonable strategy, they will witness lower productivity because of the presence of adjustment costs on newly installed capital, and a higher depreciation rate on their existing capital stock due to higher rates of capital utilization. Furthermore, under the assumption of constant returns to scale, increasing the factors in unequal proportions will reduce productivity and increase marginal cost.

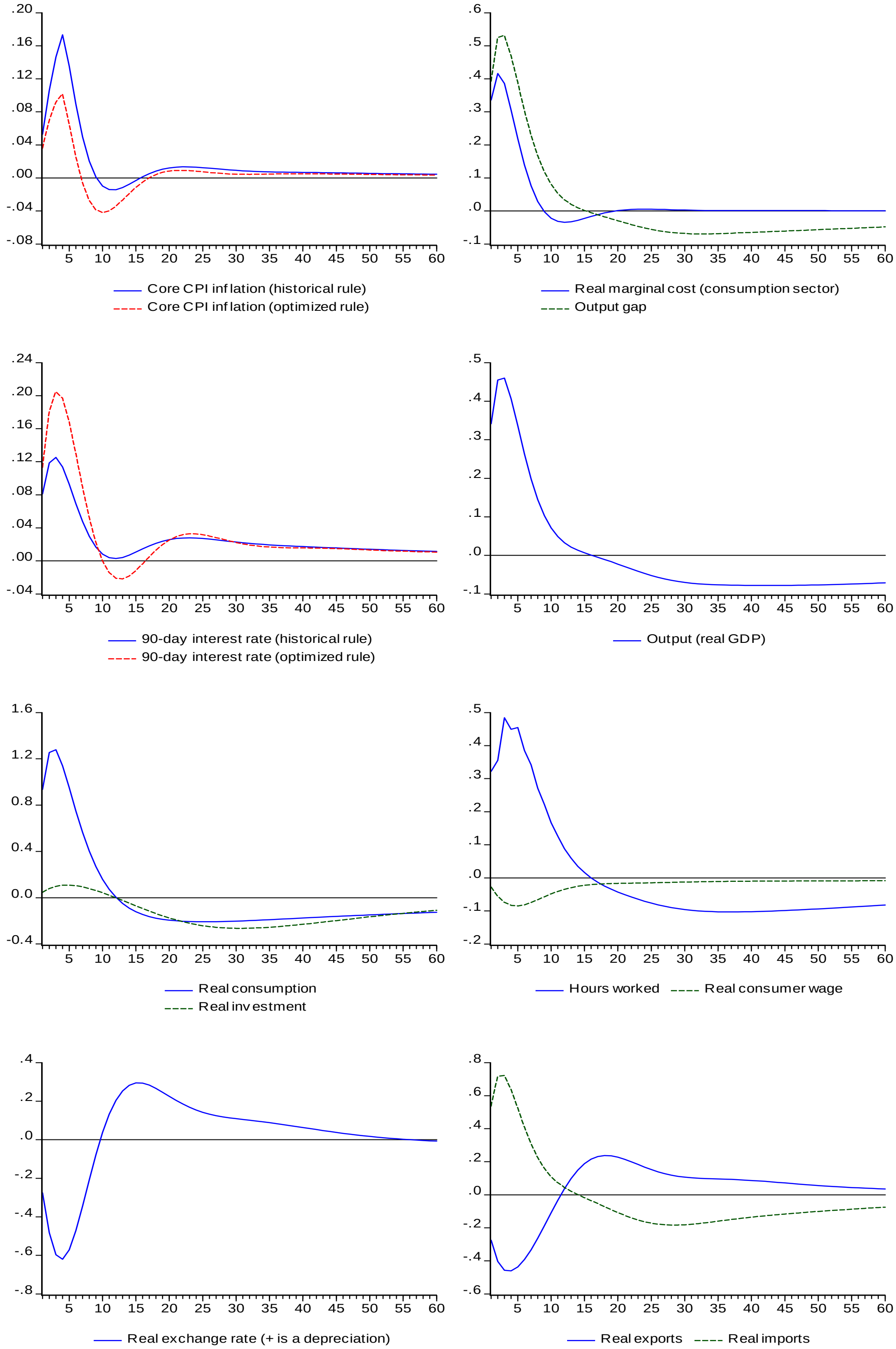
As a consequence of higher costs, all firms that are supplying more output would like to charge a higher price to maintain a constant markup. However, only a subset can change their price when the shock arrives, which means that, in aggregate, prices will rise by less than marginal cost and the average markup in the consumption-goods sector will decline.

ToTEM predicts that a 0.4 per cent increase in real marginal cost will cause year-over-year CPIX inflation to rise 0.17 basis points above target at the end of year one. It is worth mentioning that for a given increase in the output gap in ToTEM, the source of the increase is critical to the response of consumer-price inflation. For instance, the same increase in output generated by a shock to government expenditures would cause almost no change to CPIX inflation. In other words, second-round or spillover effects from one sector to another are quite small for demand-style shocks in ToTEM. Therefore, when discussing the elasticity of inflation with respect to excess demand or supply, it is important to remember that the sectoral composition matters.

So, how does monetary policy ensure such a timely return of inflation to the target in ToTEM? First, we note that the increase in expected inflation causes a modest tightening (12 basis points at its peak) by the monetary authority. However, the tightening phase lasts about 2.5 years and, in ToTEM, the duration of

**Figure 5: Consumption Shock**

Responses are based on an historically estimated rule, unless indicated



the interest rate increase is essentially as important as its size. Thus, monetary policy commits to a sustained, albeit modest, period of tighter policy that causes a rise in the real interest rate expected to prevail into the future. This increase in real interest rates, in turn, reduces the incentive of consumers to indulge in present consumption, and helps to stem the increase in investment spending in all sectors of the economy. In addition, higher rates cause a real appreciation, which facilitates a substitution towards imports, thereby alleviating some of the excess demand in the Canadian economy. The exchange rate appreciation also increases the price of Canadian exports abroad, thereby reducing export activity, further reducing excess demand pressures. All of these effects combine to help restore aggregate demand to its long-run, sustainable level, while at the same time returning inflation to its targeted level.

In the labour market, we begin by noting that the real consumer wage declines by about 0.1 per cent following the rise in consumption demand. This decline stems directly from the existence of sticky wages in ToTEM. All households would like to negotiate a higher real wage in the scenario, because they are working and consuming more. Higher consumption implies that, at the old real wage, consumption is valued relatively less than leisure, and that the only means by which this disequilibrium can be eliminated is through a higher wage or a lower supply of labour. However, as in the goods market, only a small subset of workers renegotiate their wage each period, and thus, in aggregate, the real wage falls and workers are obliged to supply whatever level of labour is demanded by firms at that wage.

As with any model that accounts for stocks, in ToTEM there is a price to be paid by consumers for their temporary spending binge. In this case, the increase in consumption is partially financed through a deterioration of the net foreign asset (NFA) position. However, because of our assumption that the desired NFA level remains unchanged in the shock, a period of dissavings must be matched by a period of increased savings, which, in ToTEM, is reflected by a sustained period of consumption that is slightly below steady state.

Regarding the distinction between the historical and the optimized rule, we first note that the optimized rule places a large weight on both the lagged interest rate term (0.95 versus 0.8) and the inflation gap (20 versus 2.5). These two differences imply that the optimized rule will be slightly more aggressive in its response to inflation deviations in the short run, and significantly more aggressive in the medium run. Of course, as we have already discussed, firms and household take into account the entire (expected) future path of interest rates when making their decisions. Therefore, the medium-term behaviour of the rule remains very important to the short-run response of the model. By credibly committing to respond more aggressively to inflation, expectations remain more closely anchored to the target, and consequently, the peak inflation response is only 60 per cent as large as with the historical rule, while nominal interest rates increase by about 20 basis points, up from 12 basis points.

### 4.4.2 *A risk-premium (exchange rate) shock*

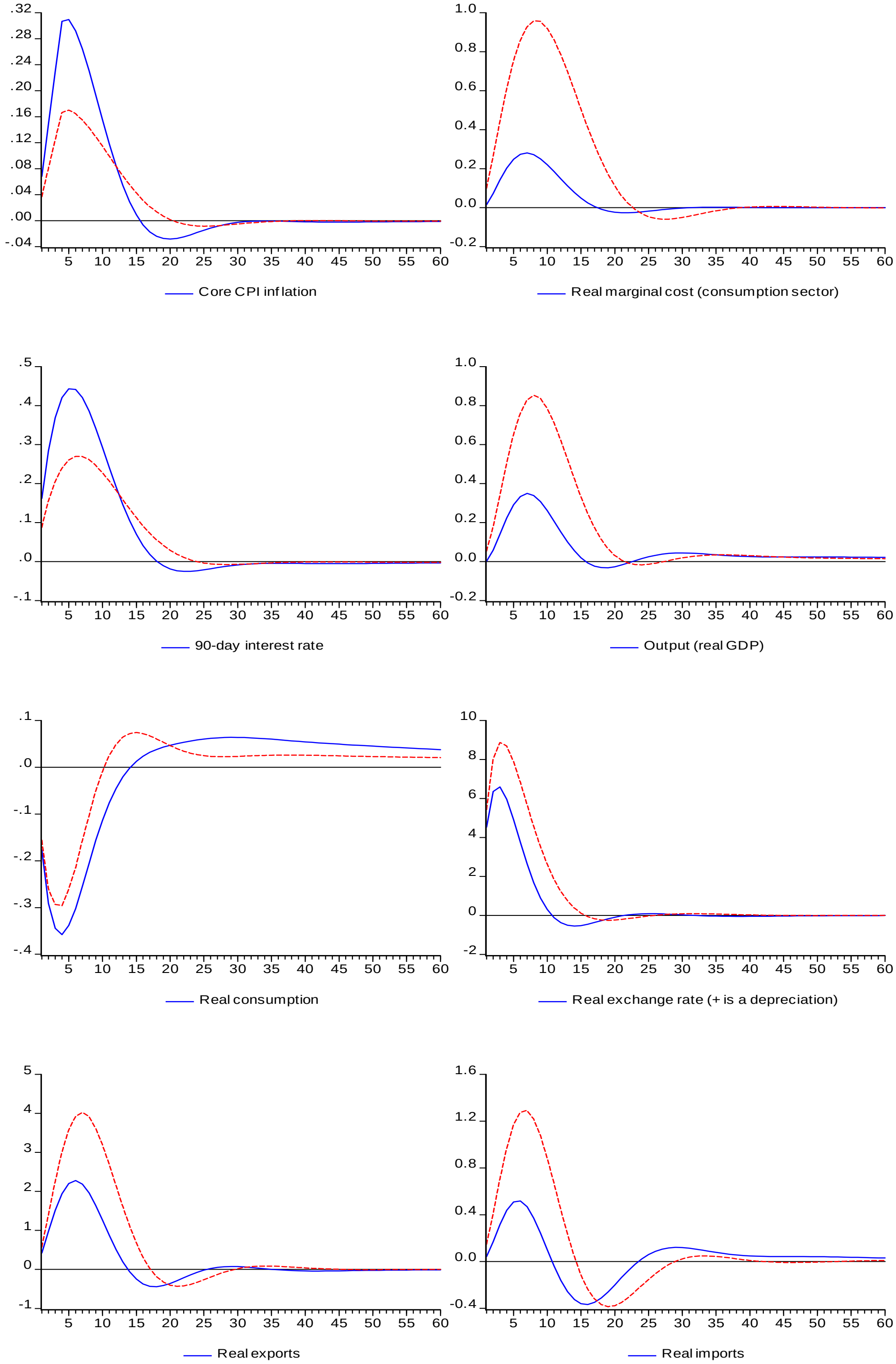
In this section, we analyze the implications of an exogenous increase to the country-specific risk premium, which has the effect of depreciating the real exchange rate by about 7 per cent by the end of the first year (Figure 6). The solid lines reflect the baseline calibration of ToTEM, which assumes a net steady-state markup of prices over a marginal cost of 5 per cent. The broken line illustrates the model's response to the same shock, assuming a markup of just over 2 per cent, which corresponds to an increase in competition in the goods market ($\epsilon$ increases from 21 to 45, and so similar goods become closer substitutes). We begin with the base-case calibration of 5 per cent.

A depreciation of the exchange rate in ToTEM causes the Canadian-dollar price of imported intermediate goods and commodities to increase, both of which are inputs to the production of final goods. Viewed from the standpoint of Canadian producers, a depreciation of the exchange rate triggered by a shock to the risk premium represents a negative supply shock. For commodity producers, who do not use imports or commodities as inputs, a depreciation is unambiguously positive, since the price they receive increases by an amount equal to the change in the exchange rate. Commodity producers respond to the price increase by expanding their production until marginal cost once again is equated with the price of their output. However, given the temporary nature of the shock, combined with the large costs associated with increasing their supply, the actual increase in commodity production is quite small.

Exporters of finished products are affected by both the supply and demand dimensions of the shock. On the one hand, the price of their inputs increases, but the demand for their output also increases. On a net basis, the depreciation causes finished-product and commodity exports to rise, while consumption falls and investment remains essentially unchanged. Unlike in QPM, imports also increase in ToTEM immediately following a depreciation of the exchange rate. This difference emerges primarily because the substitution effect is weaker in ToTEM owing to a

**Figure 6: Risk-Premium (Exchange Rate) Shock**

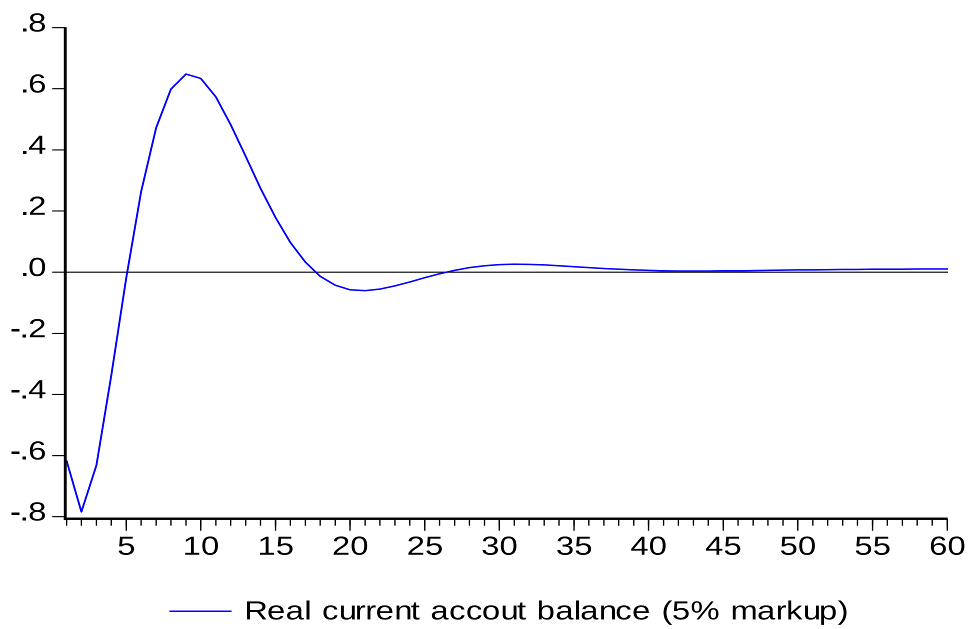
Solid line: 5% markup; broken line: 2% markup



low elasticity of substitution between imports and other factor inputs (firms are limited in how much they *can* substitute away from imports), and because the price paid by final-goods producers is quite slow to increase in response to the depreciation (firms are limited in how much they *want* to substitute). This latter effect also explains why the current account initially worsens following a depreciation. While exports rise by significantly more than imports in the short run, the price paid at the border for imports by the import distributor rises immediately by an amount equal to the change in the exchange rate (i.e., the law of one price holds at the border), whereas the Canadian-dollar price of exports is slow to adjust because of the assumption of sticky prices. Thus, an exchange rate depreciation initially causes a worsening in the terms of trade and therefore a $J$-curve effect, whereby the current account worsens for about the first year and then improves in the following years (Figure 7).

Higher input prices cause CPIX inflation to rise (to a peak of 0.3 percentage points in year two of the shock) as producers partially pass on their cost increases in the form of higher retail prices, which triggers the monetary authority to tighten policy by about 50 basis points.

**Figure 7: Current Account Response to Exchange Rate**



We turn next to the role played by competition in the model and its link to the parameters of the structural Phillips curve, the difference between the solid and broken lines in Figure 6 shows the effect of reducing the steady-state markup from 5 to 2 per cent in each of the final-product markets. When the firm-level marginal-cost curve is positively sloped, as occurs in ToTEM because of the assumption that capital is firm-owned and costly to adjust, firms are less willing to pass on higher costs to consumers. To see why, consider the impact of higher demand on those firms that are resetting their price in period $t$. These firms recognize that raising their price today will lower their market share according to the demand function:

$$\frac{C_{it}}{C_t} = \left(\frac{P_{it}^c}{P_t^c}\right)^{-\epsilon_t^c},$$

which is true regardless of the slope of the marginal-cost curve. However, when it is positively sloped, lower demand will also lower the firm's marginal cost, thereby tempering the extent to which firms *wish* to raise their price.[34] The greater is $\epsilon$ (the elasticity of substitution between finished goods in a given sector), the more competitive are markets, and the greater will be the wedge between firm-specific and average (or economy-wide) marginal cost. Thus, when markets are very competitive, demand and therefore marginal cost will be very sensitive to a firm's relative price. As a result, high competition should cause less relative-price variation and, when we aggregate the model, inflation should be less sensitive to movements in economy-wide real marginal cost, which can be confirmed by checking the Phillips curve from section 2.1:

$$\widehat{\pi}_t^c = \frac{\gamma_c}{1+\mathcal{B}\gamma}\widehat{\pi}_{t-1}^c + \frac{\mathcal{B}}{1+\mathcal{B}\gamma_c}\mathbf{E}_t\widehat{\pi}_{t+1}^c + \zeta\frac{(1-\theta_c)(1-\theta_c\mathcal{B})}{\theta_c(1+\mathcal{B}\gamma_c)}\widehat{\lambda}_t^c + \varepsilon_t^p, \tag{4.1}$$

where $\zeta < 1$.

So, by how much does increased competition reduce exchange rate pass-through in ToTEM? Recall from section 4.2 that we define pass-through as:

$$\frac{\ln(P_{t+k}/P_{t-1})}{\ln(e_t/e_{t-1})}.$$

Thus, a value of one, for example, would correspond to full pass-through. The baseline calibration predicts pass-through of about 0.09 per cent after two years,

[34] Recall that the only objective for firms is to maintain a fixed markup of price over marginal cost.

whereas, when the markup is set to 2 per cent, this value drops to just over 0.05 per cent. While the relationship is non-linear, we can say that around a base case that has a 5 per cent markup, a halving of the markup almost halves the exchange rate pass-through at a two-year horizon.

### 4.4.3 *A reduction in the inflation target*

In describing the effects of a reduction in the inflation target from 2 to 1 per cent in ToTEM, we will divide the discussion into two parts. The first concerns the size and duration of the short-run costs associated with a disinflation, whether measured in terms of foregone output, consumption, or utility, and the related issue of why these costs arise. The second issue relates to the long-run or steady-state effects on output of a lower steady-state rate of inflation on the real economy.

#### *Short-run transition costs*

In this disinflation experiment, we will first focus on the issue of the short-run costs, under the assumption that the disinflation is fully understood and credible from the viewpoint of households and firms in the model. Next, we will assess the short-run costs for the same experiment assuming that the Bank either does not benefit from full credibility, which is to say that the public does not believe that the central bank is fully committed to the policy change, or that the public does not fully recognize the nature of the policy change. This will be followed by a discussion of some of the potential long-run benefits of a lower inflation target in a sticky-price/-wage model such as ToTEM.

Figure 8 shows the response of ToTEM, under the assumption of full and partial monetary policy credibility, to a reduction in the targeted level of inflation from its current level of 2 per cent to 1 per cent. We first discuss the full-credibility case, which corresponds to the solid lines. Since there is some nominal rigidity in the economy, an unanticipated change to the target still has some effects on real variables in the short run. When prices and wages cannot fully adjust immediately, even a well-understood, fully credible disinflation will generate a small reduction in output and employment in ToTEM. So, while Figure 8 indicates that the nominal short-term interest rate never rises, the expected real interest rate does increase in the short run. The reduction in GDP (consumption) troughs at about -0.2 per cent in year two of the simulation. Owing to the reduction in investment spending over the same period, the capital stock remains below its steady state for several years, albeit by a very small amount. This, in turn, causes output to remain below its steady-state level for several years. However, the effect is temporary and output and consumption eventually return to their pre-shock levels. The increase in real interest rates also generates a fairly persistent 2 per cent appreciation of the real exchange rate, which acts to reduce net exports.

Year-over-year CPI (or CPIX) inflation takes one year to reach its new steady-state level in this simulation, and actually undershoots this level slightly in year two. Nominal short-term interest rates take about two years to converge on their new level, which is one percentage point lower, owing to lower inflation (the steady-state real interest is unchanged in this scenario). In summary, under the assumption that the disinflation is fully understood as a pure shift in the growth rate of nominal variables only, the costs in terms of output and consumption are quite small, albeit persistent.

Regarding the case of imperfect credibility, it is first useful to revisit exactly what credibility means in ToTEM. As discussed in section 2.6.1, there are two types of monetary policy disturbances in ToTEM: a change to the target, which is permanent by assumption, and a shock to the policy rule (a traditional monetary policy shock), which is transitory by assumption. Since both shocks enter the model only through the policy rule, their effects are distinguishable only insofar as their persistence differs. A target shock will cause inflation expectations to move by more than a monetary policy shock, owing to the fact that it is permanent. Imperfect credibility in ToTEM amounts to the assumption that agents cannot fully distinguish between the two shocks, and therefore must infer the source of the interest rate change, and in ToTEM this inference is done optimally using the Kalman filter. Thus, one can usefully think of the case of disinflation under imperfect credibility as being a weighted average of a shock to the target and a monetary policy shock. Indeed, under the assumption that the target has an exact unit root and the policy shock is white noise (and that the two are mutually uncorrelated), equation (2.63) reduces to:

$$\mathbf{E}_t \overline{\pi}_t = (1 - \Sigma)\, \mathbf{E}_{t-1} \overline{\pi}_{t-1} + \Sigma \overline{\pi}_t,$$

where $\overline{\pi}_t$ is the true central bank target and $\mathbf{E}_t \overline{\pi}_t$ is private agents' time $t$ perception of that target. When $\Sigma = 1$, agents learn immediately that the shock is a change to the central bank's target ($\overline{\pi}_t = \pi_t^*$ in equation (2.63)), and when $\Sigma = 0$, agents never learn, since, implicitly, agents believe that the central bank never commits to a change in the target.

The broken lines in Figure 8 correspond to the assumption that $\Sigma = 0.22$, which implies that the perceived target will equal the actual target after about 2.5 years. Under this assumption, nominal interest rates increase by about 25 basis points in the first year at the same time that actual inflation is falling. Taken together, this implies a larger and more sustained increase in real interest rates. The consequences of higher interest rates are immediately apparent from the response of consumption and investment in the model. Consumption troughs at about 0.5 per cent below control in year two, whereas investment and output trough at just over 0.5 per cent below control at the end of year two. Furthermore, the cumulative output loss under imperfect credibility after 10 years is about 2.5 per cent (implying a sacrifice ratio of 2.5), whereas under perfect credibility the loss is about 0.6 per cent of one year's output.

Finally, under imperfect credibility, it takes approximately six quarters for inflation to reach the new target, two quarters longer than with perfect credibility. This increase in time occurs despite a significantly larger marginal-cost gap in the consumption-goods sector.

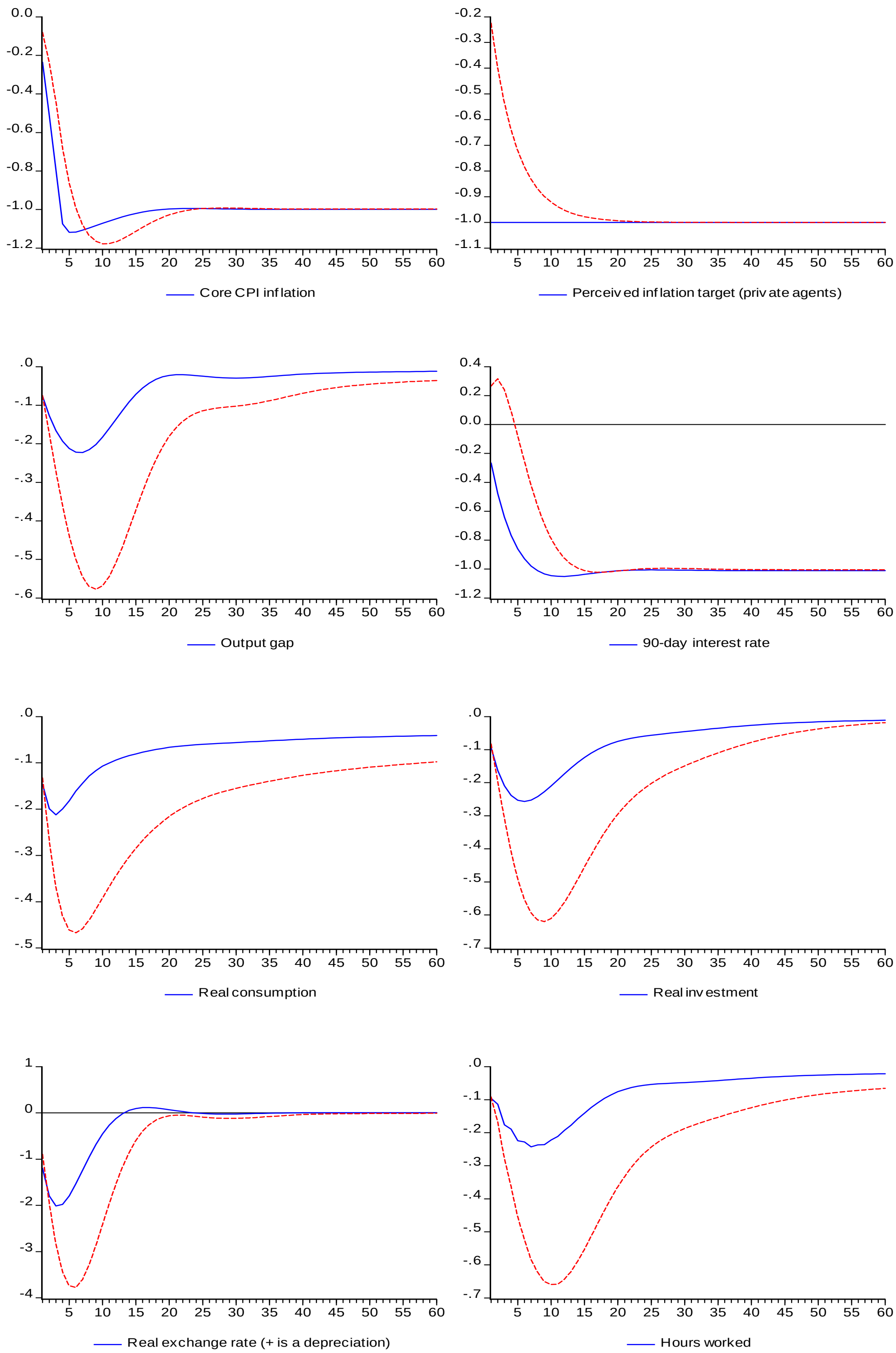
### *Long-run benefits*

Since ToTEM incorporates full indexation for prices and wages, the steady-state rate of inflation does not affect output in the model by assumption. However, this assumption is made primarily for technical reasons, not because it is viewed as reflecting the manner in which prices and wages are actually set. First, it simplifies the computation of the model's steady state and the linearization of the dynamic model in the presence of positive steady-state inflation. Second, as discussed in Hornstein and Wolman (2005), difficulties arise in aggregating the Calvo model with non-zero steady-state inflation in the presence of firm-specific factors of production such as capital, unless full indexation is assumed. Third, and most importantly, Ascari (2004) shows that the Calvo (1983) constant-hazard-function set-up greatly exaggerates the output costs of inflation precisely because it predicts that some prices (or wages) can go an arbitrarily long time without being adjusted. Therefore, while the pure Calvo set-up is convenient for most purposes, it is poorly suited to computing the steady-state effects of inflation. Consequently, for the purpose of the disinflation experiment, we use a modified version of ToTEM that incorporates the Taylor (1980) model of staggered prices and wages (Amano et al. 2006).

Positive trend inflation generates two related costs in sticky-price/-wage mod-

**Figure 8: Reduction in the Inflation Target**

Solid line: perfect credibility; broken line: imperfect credibility



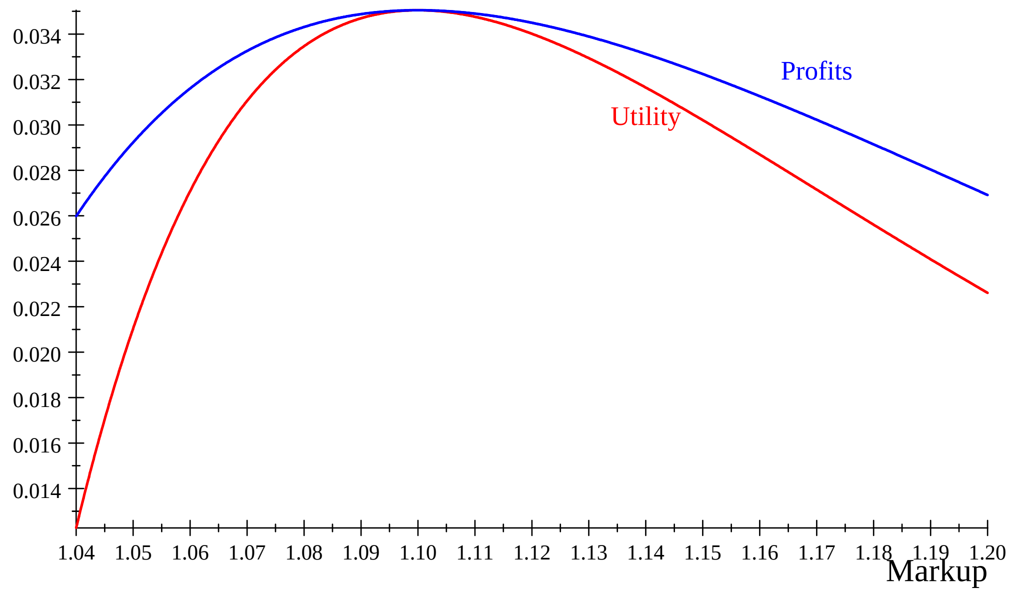
els that are directly linked to the assumption of nominal rigidity[35], meaning that money is not superneutral, even in a deterministic steady state (Ascari 2004 and Wolman 2001). The first distortion, known as the markup distortion, is due to the inherent asymmetry of the profit (and utility) function around the maximum (Figure 9).[36] In a world with positive steady-state inflation and sticky prices (wages), firms (households) can no longer achieve their desired markup in every period, because inflation erodes their relative price (wage) through time. The asymmetry implies (for a discount rate close to unity) that firms (households) will choose a higher average markup of price over marginal cost (wage over the marginal rate of substitution) when inflation is positive, because profits (utility) decline faster with a markup that is below the optimum than with a markup that is above the optimum. This higher markup is equivalent to increasing the monopoly distortion that is already present in the model due to imperfect competition. Higher markups have the effect of raising the real wage and reducing labour input (raising leisure). This, in turn, drives down the equilibrium capital stock and reduces steady-state output and consumption.

As Figure 9 shows, the asymmetry and therefore the markup distortion is larger in the labour market than in the goods market, which can be traced back to the calibrated value for $\eta$, $\eta = 0.6$. Only in the special case, $\lim_{\eta\rightarrow\infty}$, will the degree of asymmetry in the goods and labour market be equal.

[35] Several other costs may arise that are unrelated to the assumption of nominal rigidity.

[36] This asymmetry arises from the non-linearity of the demand curve for the firm's (household's) product (labour). For instance, demand for labour is given by the function $N_{ht} = \left(\frac{W_{ht}}{W_t}\right)^{-\epsilon_{w,t}} H_t$.

**Figure 9: Profits and Utility**




The second distortion is referred to as the relative-price distortion, and its effect can be seen by revisiting the CES aggregators commonly used in DSGE models. Taking the labour market aggregator as our example, we have

$$H_t \equiv \left( \int_0^1 (N_{ht})^{\frac{\epsilon_{w,t}-1}{\epsilon_{w,t}}} \, dh \right)^{\frac{\epsilon_{w,t}}{\epsilon_{w,t}-1}} . \tag{4.2}$$

For a given level of total labour input, $\int_0^1 N_{ht} dh$, (4.2) is maximized when $N_{ht} = N_{kt}$ $\vee$ $h, k$, implying that $H_t = \int_0^1 N_{ht} \, dh$. But, as already discussed, trend inflation combined with sticky wages implies a distribution of relative wages, even in steady state. Given the aggregator's demand function,

$$N_{ht} = \left( \frac{W_{ht}}{W_t} \right)^{-\epsilon_{w,t}} H_t,$$

this implies a distribution of quantities of labour demanded, meaning that $N_{ht} \neq N_{kt}$ and $H_t < \int_0^1 N_{ht} dh$. The same argument holds in the goods market.

The relative-price distortion is minimized at zero inflation, whereas the markup distortion is typically minimized at a very low positive rate of inflation.[37] Therefore, the level welfare and output are maximized in the standard sticky-price model when inflation is very close to zero. Naturally, if wages and prices are fully indexed to lagged or steady-state inflation, or some combination of the two, these distortions disappear, since all relative-price and wage dispersion is eliminated in steady state.

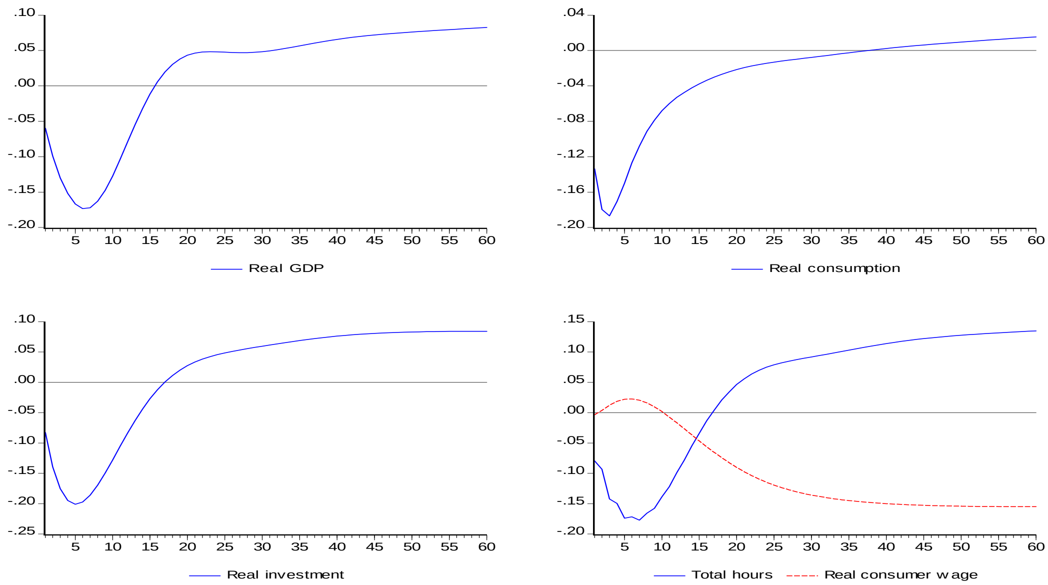
Using the calibration discussed in chapter 3, the markup distortion is about 0.5 per cent higher with 2 per cent versus 1 per cent inflation, meaning that households will demand a real wage that is 0.5 per cent higher to compensate for the effects of 2 per cent, rather than 1 per cent, inflation. The price-markup distortion, as well as the relative-price and wage distortions discussed above, while present, are quantitatively too small at this inflation rate to affect the simulation results described below.

Figure 10 shows the same full-credibility disinflation experiment, but with the steady-state labour supply effects included. As discussed above, higher trend inflation creates a disincentive to work in this class of models, or, stated otherwise, causes households to demand a higher real wage for a given level of work. Thus, when the rate of inflation is permanently reduced, labour supply increases back towards the efficient level and, as a consequence, the market-clearing real wage declines somewhat, particularly in the later years of the simulation. This positive labour-supply response means that consumption, investment, and output no longer decline by as much in the short run as in the previous scenario (about 0.2 per cent in year two). In addition, after about three years, output goes above, and remains above, its pre-shock level.

In steady state, real GDP (consumption) is slightly more than 0.1 (0.06) per cent higher. The net present benefit of the disinflation from 2 to 1 per cent under full credibility is equal to approximately 2.3 per cent of one year's output, meaning that the discounted (using $\beta$) long-run benefits exceed the short-run costs by the equivalent of 2.3 per cent of annual GDP in the economy.

[37] As the discount rate goes to one, the rate of inflation at which the markup distortion is minimized goes to zero, assuming zero technology growth in the steady state ($g = 1$). Positive technology growth will generally imply that some deflation is optimal.

**Figure 10: Reduction in the Inflation Target (including steady-state effects)**




### 4.4.4 *A commodity-price shock*

Figure 11 shows the effects in ToTEM of a persistent 10 per cent increase in the world price of commodities stemming from a supply disruption that arises in other commodity-producing countries, but that leaves Canada's commodity supply capacity unchanged. It is assumed that the shock affects the price of energy and non-energy commodities equally.

We assume that this negative supply shock has the effect of raising inflation in the rest of the world, which is countered by higher nominal interest rates and lower GDP relative to potential (a negative output gap) in these countries. Thus, from Canada's viewpoint, there are three additional exogenous changes associated with this shock, although the commodity-price increase itself is the most significant change. First, higher rest-of-world prices will raise the price of Canadian imports and make Canadian exports more competitive on world markets. Second, lower rest-of-world GDP will reduce the demand for Canadian finished-product exports. Third, higher rest-of-world interest rates will cause the Canadian dollar initially to depreciate, all else being equal.

Perhaps the most striking feature of the results is just how important commodity prices are for the Canadian economy and how long-lasting the effects of even a temporary (roughly three years) change in commodity prices can be. One of the most noteworthy effects of the increase in commodity prices is the sustained increase in consumption (about 0.4 per cent for the first five years), which lasts about 20 years. This captures the response of households to the increase in their wealth, which is captured by an immediate increase in their NFA holdings. Furthermore, given that we assume no change in households' desired NFA position, even stronger consumption is required in order to gradually restore the old NFA level.

Also noteworthy is the 2.5 per cent appreciation of the real exchange rate that builds over the first year, and then persists for several years to come.[38] This real appreciation is generated endogenously by the model, to encourage higher imports, which are needed first to stabilize the NFA position and, ultimately, to restore it to its pre-shock level. This appreciation eventually leads to a fall in the price of investment, which is import intensive, thereby boosting investment demand by as much as 0.7 per cent in year five.

On the trade side, commodity exports essentially form the residual between commodity production and demand for commodities in Canada. For a temporary shock such as this, the positive supply response by commodity producers is quite small. However, there is some substitution away from commodities by firms and consumers, owing to the price increase, and so commodity exports rise by as much as 1.4 at the end of year 1. Finished-product exports, by contrast, fall by 1.2 per cent at the end of year two, in response to the appreciation of the exchange rate and the reduction in demand in the rest of the world. Finally, imports flourish (0.5 per cent in year five), because of both an immediate and strong income effect coming from strong domestic demand, and a strong substitution effect that gradually builds through time as lower import prices at the border get passed on to manufacturing firms.

In the labour market, higher overall demand in the economy leads firms to increase their demand for labour input, measured as hours worked in ToTEM. This increase in hours, combined with stronger consumption spending, leads households to raise their desired real wage. However, since only a small subset of households can actually renegotiate their wage at the outset of the shock (recall that wage

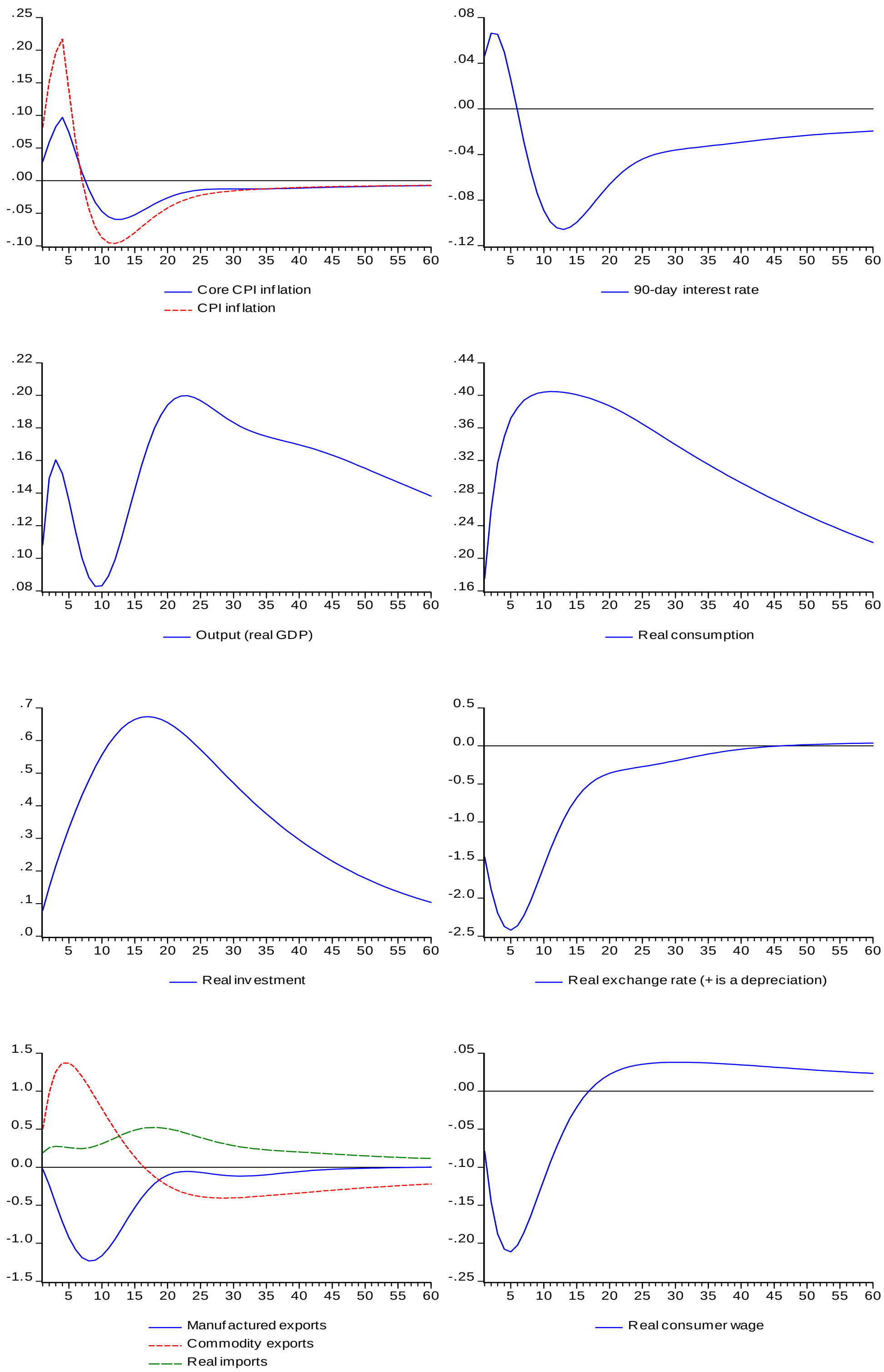
[38] If the increase in commodity prices were permanent, the real exchange rate appreciation would be more than twice as large, consistent with the cointegration evidence for Canada (Amano and van Norden 1995).

contracts are staggered and last about six quarters, on average, in ToTEM), the aggregate real consumer wage initially falls by as much as 0.1 per cent, which, from the representative firm's viewpoint, helps to stem the rise in real marginal cost. Only after about three years does the real wage rise above control.

Regarding the nominal side of the economy, CPIX inflation initially rises by as much as 0.1 per cent (0.2 per cent for CPI inflation, reflecting the higher commodity concentration in the CPI basket), and then falls below control in year three. The behaviour of inflation can be explained through the impact of commodity and import prices on real marginal cost in the consumption sector. Initially, the large increase in commodity prices, combined with an overall reduction in productivity, causes marginal cost to rise. However, as commodity prices return to control and the appreciation of the exchange rate begins to show up in the price paid by manufacturing firms for imports, real marginal cost falls below control.

Over the medium term, real GDP remains above control while inflation returns to target. This occurs because of the increase in installed capital from higher investment activity in preceding years. Thus, a persistent rise in the terms of trade in ToTEM generates a small but sustained increase in potential output. Were the commodity-price shock permanent, the rise in potential output would also be permanent and much larger in magnitude.

**Figure 11: Commodity-Price Shock**



### 4.4.5 *A technology shock*

In this experiment, we assume a 1 per cent permanent increase in the level of labour-augmenting technology, $A_t$ (see equation (2.2)), in Canada that is matched by a 1 per cent permanent increase in potential GDP in the rest of the world. In this sense, the shock may be thought of as a world productivity shock, rather than one that is unique to Canada.

A shock to the level of $A_t$ immediately raises the level of production in all sectors (Figure 12), including the commodity-producing sector.[39] In addition to higher output, this shock immediately increases the marginal productivity of all factors of production, causing firms to increase their demand for capital, commodities, and imports. As a result, a sustained investment boom occurs, as firms attempt to raise their capital stocks while at the same time minimizing the costs in terms of reduced productivity associated with installing large amounts of new capital in a short period of time. On balance, investment activity remains above its new steady-state level for approximately 15 years. Commodity and imported inputs, by contrast, attain their new steady-state level after approximately 10 years.

Unlike the other inputs to production, overall hours worked initially declines very slightly following an increase in productivity. This stems from a decline in hours in the finished-product export and government sectors. Essentially, demand is sufficiently slow to rise in these sectors (recall that prices are slow to adjust) that firms, which have become more productive, can satisfy demand with less labour. Hours worked increases in the remaining sectors.

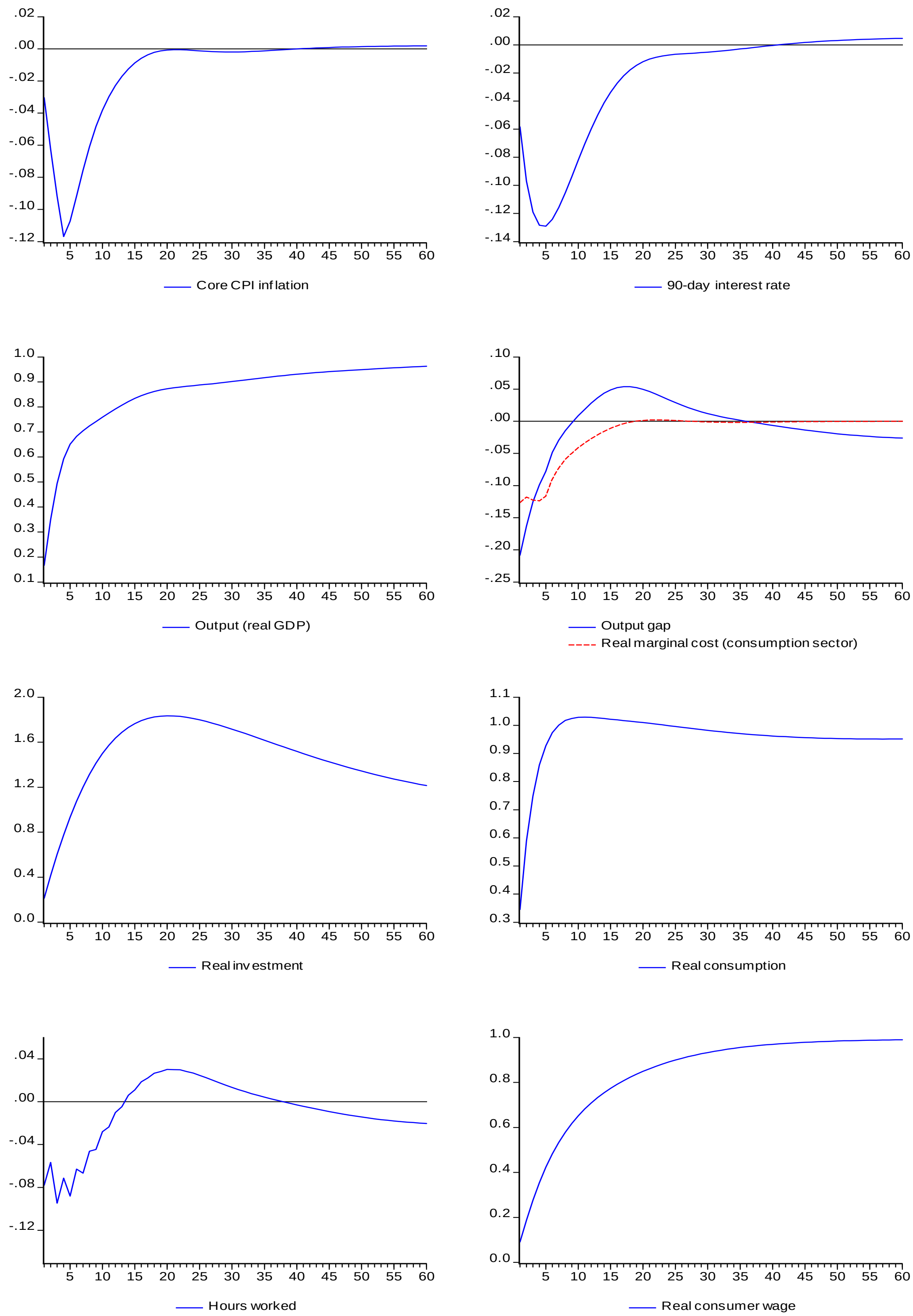
Higher productivity also causes all firms' marginal cost of production to fall, which triggers a reduction in prices throughout the economy. Focusing on the consumption sector, we see from Figure 12 that CPIX inflation troughs at about 0.12 percentage points below control at the end of the first year, which in turn causes the monetary authority to reduce nominal short-term interest rates by about 30 basis points. Interest rates remain below control for about two years, which is sufficient to generate a 1.2 per cent depreciation of the real exchange rate at the end of the first year. This depreciation helps to boost finished-product and commodity exports, and it also helps to temporarily stem the increase in demand for imported inputs.

As for the medium-term effects of higher productivity, we first note that while

[39] However, in the commodity sector, it is assumed to raise the productivity of both labour and the fixed factor land, ensuring that commodity production remains a constant proportion of GDP, independent of the level of $A_t$.

the adjustment process takes several years, all of the components of national income eventually rise by 1 per cent (the size of the shock to $A$), as does GDP, meaning that ratios such as consumption divided by GDP remain unchanged following an increase in world productivity. On the supply side, all inputs to production except labour input also rise by 1 per cent, reflecting our choice of utility function. The fact that labour supply does not increase means that the shock to technology is fully absorbed by the real producer wage, which also eventually increases by 1 per cent. In addition, the fact that the capital-to-output ratio remains unchanged in the long run reflects our choice of technology in the production process. By making technology ($A_t$) labour-augmenting only, the capital-to-output ratio remains stationary in the presence of trend technology growth.

**Figure 12: Technology Shock**



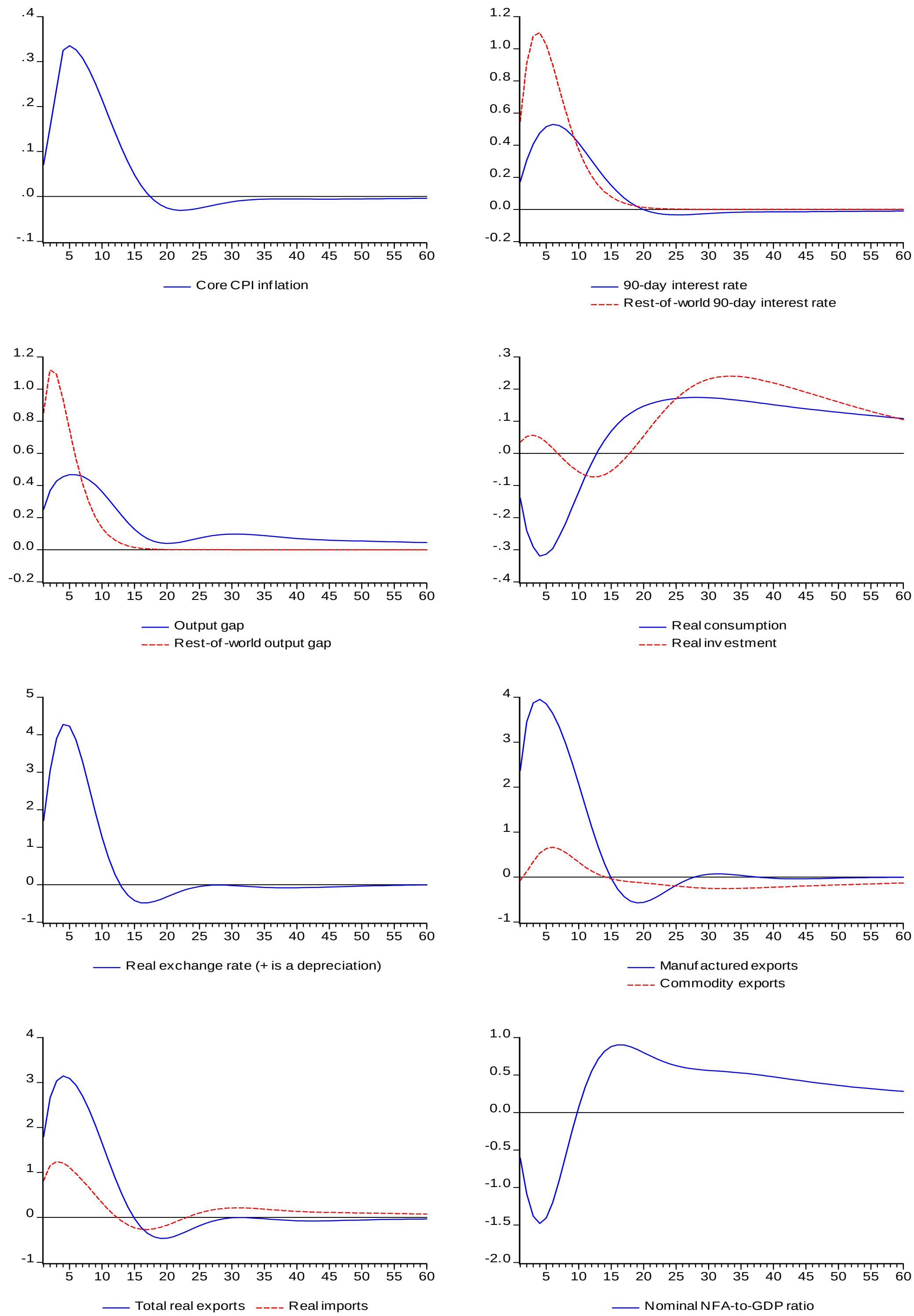
### 4.4.6 *A rest-of-world demand shock*

In this scenario, we consider the implications of a shock to aggregate demand among Canada's major trading partners that causes a peak increase in the trade-weighted output gap of about 1.1 percentage points (Figure 13). The responses of the other relevant rest-of-world variables have been calibrated to match the response of the Bank of Canada's model of the U.S. economy, MUSE (Gosselin and Lalonde 2005); in this sense, one can think of this shock as one that arises in the United States only.

The increase in the trade-weighted output gap is fairly persistent, lasting about three years. As a result, the rest-of-world inflation and interest rate responses are also quite persistent. Strong foreign demand and higher foreign inflation and interest rates cause an increase in the demand for Canada's finished-product exports, and a depreciation of the exchange rate. Higher demand for Canadian exports arises not only because of higher rest-of-world demand, but also because of the combination of higher prices abroad and a weaker Canadian dollar, both of which make Canadian exports more competitive on world markets. Taken together, total exports rise just over 3 per cent in year two of the shock, which is sufficient to raise real GDP in Canada by slightly less than 0.5 per cent near the end of year two, despite modest declines in consumption and investment spending over the same period.

Higher prices for imported intermediate goods, investment goods, and commodities, all driven by the depreciation of the exchange rate, cause real marginal cost to rise in the consumption sector, which in turn puts upward pressure on CPIX inflation (CPI inflation rises by somewhat more, reflecting its higher commodity component). Year-over-year CPIX inflation increases to just over 0.3 percentage points above control in year two, which prompts the monetary authority to raise nominal short-term interest rates by as much as 50 basis points, about half of the increase seen abroad.

**Figure 13: Rest-of-World Demand Shock**



### 4.4.7 *A price-markup shock*

Interpreted literally, a price shock in ToTEM represents a temporary increase in the desired markup in the consumption-goods sector (see equation (2.21)). More generally, it can be interpreted as a negative supply/price shock to one of the components of the CPIX basket, which has the effect of increasing the quarterly rate of CPIX inflation for about one year. The shock is quite large (Figure 14), causing year-over-year CPIX inflation to increase by just over 1 percentage point at the end of the first year. Given its persistent nature, monetary policy responds by raising nominal interest rates by about 50 basis points and then running this increase off over the course of about one year.

The combination of higher prices for consumption goods (in terms of the wage) and higher real interest rates acts to reduce consumption demand both in the short run and in the medium run. The initial reduction in aggregate consumption reflects a large decline in consumption by current-income households, whose spending falls one-for-one with the reduction in their real wage, and a reaction of lifetime-income consumers to a combination of higher real interest rates and slightly lower lifetime income. As the real wage is restored to its pre-shock level, current-income consumption returns to normal, whereas lifetime-consumption households are still consuming less (recall that lifetime-income consumers smooth their consumption to the extent possible, and therefore adjust their consumption by small amounts over an extended period of time in the face of shocks such as this).

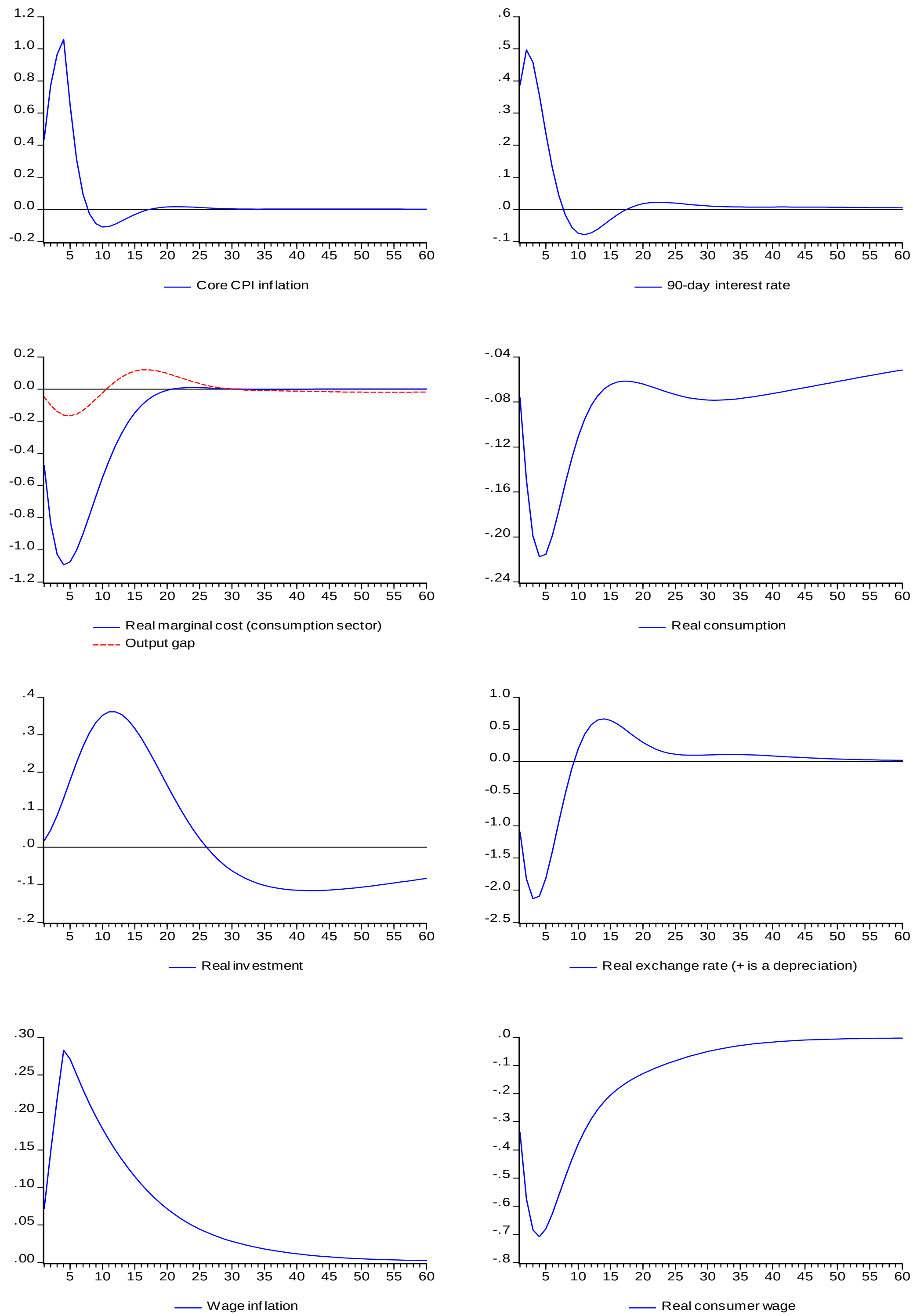
The combination of lower consumption and export activity is sufficient to reduce output (and the output gap) by about 0.15 per cent below control at the end of the first year. However, output then rises above control, beginning in year three, reflecting an increase in demand for finished-product exports. This effect is driven primarily by the fact that the nominal exchange rate (whose response is initially dominated by the interest-rate increase) ultimately adjusts to the higher price level faster than does the price of finished-product exports, which is sticky in the short run.[40] Thus, by the beginning of year three, the rest-of-world dollar price of Canadian finished-products has declined, which raises demand.

Perhaps surprisingly, aggregate investment rises in this scenario, albeit with a reasonable lag, reflecting higher investment activity in all but the consumption sector. This mainly reflects a decline in the relative price of investment, stemming from a reduction in the relative price of imports, a key input to the production of

[40] Recall that purchasing power parity holds in this scenario, which means that the exchange rate will ultimately increase by the overall price level.

manufactured investment goods. Furthermore, while nominal interest rates rise in the scenario, they do so in response to higher inflation; when viewed in terms of the expected real interest rate, the increase is quite modest.

**Figure 14: Markup (Price) Shock**



# Chapter 5

# Concluding Remarks

The decision to embark on building a new model for projections and policy analysis was based on several considerations. First and foremost, there was a strong sense among Bank staff that more structure could now be incorporated into a macro model without sacrificing the model's ability to explain the historical data relative to QPM, and that this added structure would be helpful for understanding, and predicting the implications of, shocks to the Canadian economy. In other words, while staff continue to believe that a trade-off exists between structure and data fit, the extent of this trade-off has diminished with the introduction of recent DSGE models, such as that described in CEE. Along this margin, we judge ToTEM to be a success. ToTEM provides a richer and more complete interpretation of a wider array of shocks, which makes it a more useful tool for both the projection and for policy analysis. Furthermore, informal comparisons suggest that ToTEM provides a better fit to the data than QPM.

The second objective was to have a model that would be easier to use, modify, and learn by staff. By adopting a model the core of which is essentially that of an open-economy DSGE model, recent university graduates hired by the Bank will already be familiar with the basic structure of ToTEM, and therefore less training will be required. In addition, the basic optimizing-agent structure underlying ToTEM is very flexible. Additional features developed in the academic literature or at other central banks can be introduced (or turned off) in ToTEM in a much more straightforward manner than with QPM. Finally, the use of linearization and new solution techniques allow the staff to simulate ToTEM in a fraction of the time required with QPM.[41] These techniques also mean that the staff can more easily

[41] The kinked Phillips curve in QPM precluded the use of linear approximations.

compute ToTEM's historical covariance matrix of shocks, which was very difficult in QPM; therefore, higher moments in ToTEM, such as variances and covariances, can be computed much more quickly and precisely. This, in turn, will make model evaluation exercises along the lines of Amano et al. (2002) easier to conduct.

That said, no model is ever complete or entirely satisfactory. Indeed, the decision to be an early adopter of this new class of models—the Bank of Canada is one of the first central banks to use a DSGE model to produce the main projection—means that ToTEM, much like QPM in the early 1990s, will experience some growing pains. The staff expect to learn a lot about the model's real-world projection properties over the next year, and this experience will likely lead to certain modifications.

With respect to planned model development, the work can be roughly divided into a set of smaller, near-term projects and a set of more substantial, medium-term projects. In the shorter run, the staff plan to revisit the supply side of the model. Currently, each type of firm combines capital services, labour, commodities, and imports to produce a finished good, but only the relative import concentration distinguishes these goods in steady state. Future versions will allow for differences in the relative intensities of all factor inputs. Work is also planned to better capture adjustment costs in the production process, make an explicit distinction between energy and non-energy commodities, introduce a role for inventories, and allow a separate role for average hours and employment in the model.

Work is currently under way to formally estimate ToTEM and to conduct a comprehensive evaluation of the model's empirical properties.[42] In general, the benefits of formal estimation over our current approach of informal calibration are twofold. First, it should help the model to make more accurate forecasts, which is clearly a benefit given its intended use. Second, formal estimation would yield a measure of the uncertainty associated with the parameter estimates that could be used to assess risks to the projection, construct confidence intervals, and aid in the design of more robust monetary policy rules. However, as discussed in section 3.1, determining the most appropriate way to estimate a DSGE model is not a straightforward task. Early work that focused on matching impulse-response functions from a VAR drew criticism because the identifying restrictions used for the VAR did not hold in the DSGE model. One way to circumvent this problem would be to impose a set of just-identifying restrictions from ToTEM on the VAR. Alternatively, we could impose restrictions on the information structure (the dating

[42] A second Bank of Canada technical report will provide a comprehensive evaluation of ToTEM's empirical properties, along with a set of formal parameter estimates.

of expectations) in ToTEM, to replicate the identification structure used in the VAR. Another approach altogether would be to adopt Bayesian techniques, as in Smets and Wouters (2005), and estimate the model using maximum likelihood, although the source of these priors in many cases remains unclear.

Finally, in the medium term, the staff plan to revisit the manner in which expectations are formed in ToTEM. While appropriate in most cases, purely rational expectations can be unrealistic under certain circumstances, particularly when unusual shocks that are not well understood by private agents hit the economy. Future versions of ToTEM will offer greater flexibility in the treatment of expectations. Work is also under way at the Bank to introduce a financial sector into a small DSGE model (Christensen et al. 2006). Once completed, the staff plan to examine carefully the potential benefits of introducing such a sector in ToTEM.

While we stress that ToTEM is more structural or has better microeconomic foundations than QPM, this should not be taken to mean that the staff regard ToTEM as micro-founded in an absolute sense. Indeed, there are aspects of ToTEM, and other recent DSGE models, where this term is clearly inappropriate. For instance, no truly micro-founded model would treat the timing of price adjustment as being beyond the control of the firm, and while it is often asserted that this class of models stands up well to the Lucas critique, there is scarcely little evidence in the literature to support this assertion's main implication, that the model's structural parameters be stable across different monetary policy regimes. We do assert, however, that DSGE models represent a big improvement along this margin relative to previous-generation models. Academics and central bank economists will continue to find innovative ways to introduce more realistic microfoundations in tractable ways, and policy models such as ToTEM will no doubt benefit from these innovations.

# References


Altig, D., L. Christiano, M. Eichenbaum, and J. Linde. 2004. "Firm Specific Capital, Nominal Rigidities and the Business Cycle." Manuscript, Northwestern University.

Amano, R., D. Coletti, and T. Macklem. 1999. "Monetary Rules When Economic Behaviour Changes." Bank of Canada Working Paper No. 99–8.

Amano, R., K. McPhail, H. Pioro, and A. Rennison. 2002. "Evaluating the Quarterly Projection Model: A Preliminary Investigation." Bank of Canada Working Paper No. 2002–20.

Amano, R., K. Moran, S. Murchison, and A. Rennison. 2006. "On the Long Run Benefits of Price Stability for Canada." Manuscript, Bank of Canada.

Amano, R. and S. Murchison. 2005. "Factor-Market Structure, Shifting Inflation Targets, and the New Keynesian Phillips Curve." In *Issues in Inflation Targeting*. Proceedings of a conference held by the Bank of Canada, April 2005. Ottawa: Bank of Canada.

Amano, R. and S. van Norden. 1995. "Terms of Trade and Real Exchange Rates: The Canadian Evidence." *Journal of International Money and Finance* 14(1): 83–104.

Amirault, D., C. Kwan, and G. Wilkinson. 2006. "Survey of Price-Setting Behaviour of Canadian Companies." Bank of Canada Working Paper No. 2006–35.

Anderson, G. 1997. "Continuous Time Application of the Anderson-Moore (AIM) Algorithm for Imposing the Saddle Point Property in Dynamic Models." Board of Governors of the Federal Reserve System. Unpublished manuscript.

Anderson, G. and G. Moore. 1985. "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models." *Economics Letters* 17: 247–52.

Andolfatto, D., S. Hendry, and K. Moran. 2002. "Inflation Expectations and Learning about Monetary Policy." Bank of Canada Working Paper No. 2002–30.

Ascari, G. 2004. "Staggered Prices and Trend Inflation: Some Nuisances." *Review of Economic Dynamics:* 7. 642–67.

Ball, L. 1994. "Credible Disinflation with Staggered Price Setting." *American Economic Review* 84. 282–89.

Batini, N. and E. Nelson. 2001. "Optimal Horizons for Inflation Targeting." *Journal of Economic Dynamics and Control* 25: 891–910.

Black, R., D. Coletti, and S. Monnier. 1997. "On the Costs and Benefits of Price Stability." In *Price Stability, Inflation Targets and Monetary Policy*, 303–42. Proceedings of a conference held by the Bank of Canada, May 1997. Ottawa: Bank of Canada.

Butler, L. 1996. *The Bank of Canada's New Quarterly Projection Model Part 4. A Semi-Structural Method to Estimate Potential Output: Combining Economic Theory with a Time-Series Filter.* Technical Report No. 77. Ottawa: Bank of Canada.

Calvo, G. 1983. "Staggered Contracts and Exchange Rate Policy." *Exchange Rates and International Economics*, edited by J. Frenkel. Chicago: University of Chicago Press.

Cayen, J.P., A. Corbett, and P. Perrier. 2006. "An Optimized Monetary Policy Rule for ToTEM." Bank of Canada Working Paper No. 2006–41.

Christensen, I., B. Fung, and C. Meh. 2006. "Modelling Financial Channels for Monetary Policy Analysis." *Bank of Canada Review* (Autumn): 33–40.

Christiano, L., M. Eichenbaum, and C. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 133(1): 1–45.

Coletti, D., B. Hunt, D. Rose, and R. Tetlow. 1996. *The Bank of Canada's New Quarterly Projection Model Part 3: The Dynamic Model: QPM.* Technical Report No. 75. Ottawa: Bank of Canada.

Coletti, D. and S. Murchison. 2002. “Models in Policy Making.” *Bank of Canada Review* (Summer): 19–26.

Coletti, D., J. Selody, and C. Wilkins. 2006. “Another Look at the Inflation-Target Horizon.” *Bank of Canada Review* (Summer): 31–37.

Demers, F. 2003. “The Canadian Phillips Curve and Regime Shifting.” Bank of Canada Working Paper No. 2003–32.

Duguay, P. 1994. “Empirical Evidence on the Strength of the Monetary Transmission Mechanism in Canada: An Aggregate Approach.” *Journal of Monetary Economics* 33: 39–61.

Duguay, P. and D. Longworth. 1998. “Macroeconomic Models and Policy-Making at the Bank of Canada.” *Economic Modelling* 15: 357–75.

Eichenbaum, M. and J. Fisher. 2004. “Evaluating the Calvo Model of Sticky Prices.” NBER Working Paper No. 10617.

Erceg, C.J., D.W. Henderson, and A.T. Levin. 2000. “Optimal Monetary Policy with Staggered Wage and Price Contracts.” *Journal of Monetary Economics* 46: 281–313.

Erceg, C.J. and A.T. Levin. 2003. “Imperfect Credibility and Inflation Persistence.” *Journal of Monetary Economics* 50: 915–44.

Gagnon, E. and H. Khan. 2005. “New Phillips Curve Under Alternative Production Technologies for Canada, the United States and the Euro Area.” *European Economic Review* 49: 1571–1602.

Goodfriend, M. and R. King. 1997. “The New Neoclassical Synthesis and the Role of Monetary Policy.” In *NBER Macroeconomics Annual 1997*, edited by B. Bernanke and J. Rotemberg. Cambridge, Mass.: MIT Press.

Gosselin, M.A., and R. Lalonde. 2005. *MUSE: The Bank of Canada’s New Projection Model of the U.S. Economy.* Technical Report No. 96. Ottawa: Bank of Canada.

Hornstein, A. and A. Wolman. 2005. “Trend Inflation, Firm-Specific Capital, and Sticky Prices.” Federal Reserve Bank of Richmond *Economic Quarterly* Volume 91/4 (Fall).

International Monetary Fund. 2004. "GEM: A New International Macroeconomic Model." IMF Occasional Paper No. 239.

Jenkins, P. and D. Longworth. 2002. "Monetary Policy and Uncertainty." *Bank of Canada Review* (Summer) 3–10.

King, R.G., C.I. Plosser, and S.T. Rebelo. 1988. "Production, Growth and Business Cycles: 1. The Basic Neoclassical Model." *Journal of Monetary Economics* 21: 195–232.

Lam, J.P. and G. Tkacz. 2004. "Estimating Policy-Neutral Interest Rates for Canada Using a Dynamic Stochastic General-Equilibrium Framework." Bank of Canada Working Paper No. 2004–09.

Lane, P. 2001. "The New Open Economy Macroeconomics: A Survey." *Journal of International Economics*, (August) 54(2): 235–66.

Levin, A., F. Natalucci, and J. Piger. 2004. "Explicit Inflation Objectives and Macroeconomic Outcomes." European Central Bank Working Paper No. 383.

Levin, A. and J. Piger. 2004. "Is Inflation Persistence Intrinsic in Industrial Economies?" European Central Bank Working Paper No. 334.

Longworth, D. 2002. "Inflation and the Macroeconomy: Changes from the 1980s to the 1990s." *Bank of Canada Review* (Spring): 3–18.

Lucas, Jr., R.E. 1976. "Econometric Policy Evaluation: A Critique." In *The Phillips Curve and Labor Markets. Carnegie-Rochester Conference on Public Policy* 1: 19–46. New York: North-Holland.

Macklem, T. 1992. "Terms-of-Trade Shocks, Real Exchange Rate Adjustment, and Sectoral and Aggregate Dynamics." In *The Exchange Rate and the Economy*, 1–60. Proceedings of a conference held at the Bank of Canada, 22–23 June 1992. Ottawa: Bank of Canada.

——— 1993. "Terms-of-Trade Disturbances and Fiscal Policy in a Small Open Economy." *Economic Journal* 103: 916–36.

——— 2001. "A New Measure of Core Inflation." *Bank of Canada Review* (Autumn): 3–12.

——— 2002. "Information and Analysis for Monetary Policy: Coming to a Decision." *Bank of Canada Review* (Summer): 11–18.

Macklem, T., D. Rose, and R. Tetlow. 1995. "Government Debt and Deficits in Canada: A Macro Simulation Analysis." Bank of Canada Working Paper No. 95–4.

Maclean, D. and H. Pioro. 2000. "Price-Level Targeting–The Role of Credibility." In *Price Stability and the Long-Run Target for Monetary Policy.* Proceedings of a conference held by the Bank of Canada, June 2000. Ottawa: Bank of Canada.

McCallum, B. and E. Nelson. 2000. "Monetary Policy for an Open Economy: An Alternative Framework with Optimizing Agents and Sticky Prices." *Oxford Review of Economic Policy* 16: 74–91.

Murchison, S. 2001. "NAOMI: A New Quarterly Forecasting Model: Part II A Guide to Canadian NAOMI." Department of Finance Canada Working Paper No. 2001-25.

Murchison, S., A. Rennison, and Z. Zhu. 2004. "A Structural Small Open-Economy Model for Canada." Bank of Canada Working Paper No. 2004–04.

Perrier, P. 2005. "La fonction de production et les données canadiennes." Bank of Canada Working Paper No. 2005–20.

Plourde, A. 1987. "The Impact of $(US)15 Oil on the Canadian Economy: Evidence from the MACE Model." *Canadian Public Policy* 13(1): 19–25.

Schmitt-Grohé, S. and M. Uribe. 2003. "Closing Small Open Economy Models." *Journal of International Economics* 61(1): 163–85.

Selody, J. 2001. "Uncertainty and Multiple Perspectives." In *Monetary Analysis: Tools and Applications,* edited by H.-J. Klöckers and C. Willeke, 31–45. Frankfurt: European Central Bank.

Smets, F. and R. Wouters. 2005. "Comparing Shocks and Frictions in US and Euro Area Business Cycles: A Bayesian DSGE Approach." *Journal of Applied Econometrics* 20(2): 161–83.

————. 2002. "Openness, Imperfect Exchange Rate Pass-through and Monetary Policy." *Journal of Monetary Economics* 49: 947–81.

Taylor, J. 1980. "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88: 1–23.

Wolman, A. 1999. "Sticky Prices, Marginal Cost, and Behavior of Inflation." *Federal Reserve Bank of Richmond Economic Quarterly* 85(4): 29–48.

————. 2001. "A Primer on Optimal Monetary Policy with Staggered Price-Setting." *Federal Reserve Bank of Richmond Economic Quarterly* 87 (Fall) 27–52.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton, NJ: Princeton University Press.

————. 2005. "Firm-Specific Capital and the New-Keynesian Phillips Curve." *International Journal of Central Banking* 1(2): 1–46.

# Appendix A

# The Non-Linear Model

## A.1 Finished-products sector

As described in section 2.1, net output of the core consumption good is given by

$$\begin{aligned} C_{it} &= \mathcal{F}\left(\mathcal{H}, M_{it}^c\right) - \frac{\chi_k}{2}\left(\frac{I_{it}^c}{K_{it}^c} - (\omega + g - 1)\right)^2 K_{it}^c - \frac{\chi_I}{2}\left(\frac{I_{it}^c}{I_{i,t-1}^c} - g\right)^2 I_t \\ &\quad -\rho\left(1 - e^{\chi_u\left(u_{it}^c - 1\right)}\right)\frac{P_t^I}{P_t^c}K_{it}^c - \frac{\chi_h}{2}\left(\frac{H_t^c}{H_{t-2}^c} - 1\right)^2 A_t, \end{aligned} \tag{A.1}$$

the Lagrangian of which will be represented by $\lambda_{it}$. The production technology, which we write here as $\mathcal{F}\left(L_{it}^c, M_{it}^c, K_{i,t}^c, u_{it}^c, COM_{it}^c\right)$, is a nested CES function in five choice variables for the $i$th firm. In addition to the production technology, the firm faces constraints imposed by the capital accumulation equation, whose associated Lagrangian will be represented by $q_{it}$, and the constarint equating supply with demand, whose associated Lagrangian will be represented by $\lambda_{it}$. Thus, the firm's problem is to choose $C_{it}$, $L_{it}^c$, $K_{i,t+1}^c$, $I_{it}^c$, $u_{it}^c$, $COM_{it}^c$, and $M_{it}^c$ to maximize in expectation the following:

$$\mathcal{L}_t = \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s} \left\{ \begin{array}{c} P_s^c C_{is} - W_s H_{is}^c - P_s^{com} COM_{is}^c - P_s^I I_{is}^c - P_s^m M_{is}^c \\ -q_{is}^c \left( K_{i,s+1}^c - (1-\omega) K_{is}^c - I_{is}^c \right) \\ +\lambda_{is} \left( \begin{array}{c} -C_{is} + \mathcal{F}\left( L_{is}^c, M_{is}^c, K_{is}^c, u_{is}^c, COM_{is}^c \right) \\ -\frac{\chi_K}{2} \left( \frac{I_{is}^c}{K_{is}^c} - (\omega + g - 1) \right)^2 K_{is}^c \\ -\frac{\chi_I}{2} \left( \frac{I_{is}^c}{I_{i,s-1}^c} - g \right)^2 I_s - \rho \left( 1 - e^{\chi_u (u_{is}^c - 1)} \right) \frac{P_s^I}{P_s^c} K_{is}^c \\ -\frac{\chi_h}{2} \left( \frac{H_s^c}{H_{s-2}^c} - 1 \right)^2 A_s \end{array} \right) \end{array} \right\}. \tag{A.2}$$

Optimal investment will be obtained by setting $\frac{\partial \mathcal{L}_t}{\partial K_{i,t+1}^c} = \frac{\partial \mathcal{L}_t}{\partial I_{it}^c} = 0$ with

$$\begin{aligned} \frac{\partial \mathcal{L}_t}{\partial K_{i,t+1}^c} &= 0 = -q_{it}^c + \mathcal{R}_{t,t+1}(1-\omega) q_{i,t+1}^c \\ &\quad + \mathcal{R}_{t,t+1} \lambda_{i,t+1} \left( \begin{array}{c} \frac{\partial \mathcal{F}(\cdot)}{\partial K_{i,t+1}^c} - \frac{\chi_K}{2} \left( \frac{I_{i,t+1}^c}{K_{i,t+1}^c} - (\omega + g - 1) \right)^2 \\ + \chi_K \left( \left( \frac{I_{i,t+1}^c}{K_{i,t+1}^c} - (\omega + g - 1) \right) \frac{I_{i,t+1}^c}{K_{i,t+1}^c} \right) \\ -\rho \left( 1 - e^{\chi_u (u_{i,t+1}^c - 1)} \right) \frac{P_{t+1}^I}{P_{t+1}^c} \end{array} \right), \\ q_{it}^c &= (1-\omega) \frac{q_{i,t+1}^c}{1 + R_t} \\ &\quad + \frac{\lambda_{i,t+1}}{1 + R_t} \left( \begin{array}{c} \frac{\partial \mathcal{F}(\cdot)}{\partial K_{i,t+1}^c} + \frac{\chi_K}{2} \left( \left( \frac{I_{i,t+1}^c}{K_{i,t+1}^c} \right)^2 - (\omega + g - 1)^2 \right) \\ -\rho \left( 1 - e^{\chi_u (u_{i,t+1}^c - 1)} \right) \frac{P_{t+1}^I}{P_{t+1}^c} \end{array} \right), \end{aligned} \tag{A.3}$$

$$\begin{aligned} \frac{\partial \mathcal{L}_t}{\partial I_{it}^c} &= 0 = -P_t^I + q_{it}^c \\ &\quad - \lambda_{it} \left( \begin{array}{c} \chi_K \left( \frac{I_{it}^c}{K_{it}^c} - (\omega + g - 1) \right) + \chi_I \left( \frac{I_{it}^c}{I_{i,t-1}^c} - g \right) \frac{I_{it}^c}{I_{i,t-1}^c} \\ + \frac{\chi_I}{2} \left( \frac{I_{it}^c}{I_{i,t-1}^c} - g \right)^2 \end{array} \right) \\ &\quad + \mathcal{R}_{t,t+1} \lambda_{i,t+1} \chi_I \left( \frac{I_{i,t+1}^c}{I_{it}^c} - g \right) \left( \frac{I_{i,t+1}^c}{I_{it}^c} \right)^2. \end{aligned} \tag{A.4}$$

Next, we solve for the demand for hours by setting $\frac{\partial \mathcal{L}_t}{\partial H_t^c} = 0$:

$$\begin{aligned}
\frac{\partial \mathcal{L}_t}{\partial H_t^c} &= -W_t + \lambda_{it}\left(\frac{\partial \mathcal{F}(\cdot)}{\partial H_t^c} - \chi_h A_t\left(\frac{H_t^c}{H_{t-2}^c} - 1\right)\frac{1}{H_{t-2}^c}\right) \\
&\quad + \mathcal{R}_{t,t+2}\lambda_{i,t+2}\chi_h A_{t+2}\left(\left(\frac{H_{t+2}^c}{H_t^c} - 1\right)\frac{H_{t+2}^c}{\left(H_t^c\right)^2}\right) \\
&= 0.
\end{aligned}$$

Capital utilization is given by

$$\frac{\partial \mathcal{L}_t}{\partial u_t^c} = \lambda_{it}\left(\frac{\partial \mathcal{F}(\cdot)}{\partial u_{it}^c} + \rho\chi_u e^{\chi_u(u_{it}^c - 1)}\frac{P_t^I}{P_t^c}K_{it}^c\right) = 0.$$

The remaining demand functions are given by

$$\begin{aligned}
\frac{\partial \mathcal{F}(\cdot)}{\partial COM_{it}^c} &= \frac{P_t^{com}}{\lambda_{it}} \\
\frac{\partial \mathcal{F}(\cdot)}{\partial L_{it}^c} &= \frac{W_t}{\lambda_{it}} \\
\frac{\partial \mathcal{F}(\cdot)}{\partial M_{it}^c} &= \frac{P_t^m}{\lambda_{it}},
\end{aligned}$$

where, for instance, $\frac{\partial \mathcal{F}(\cdot)}{\partial L_{it}^c} = \frac{\partial \mathcal{F}(\cdot)}{\partial \mathcal{H}(\cdot)}\frac{\partial \mathcal{H}(\cdot)}{\partial \mathcal{G}(\cdot)}\frac{\partial \mathcal{G}(\cdot)}{\partial L_{it}^c}$, and so on.

Finally, to obtain the first-order condition for contract prices, first note that the probability in period $s$ that a price chosen in period $t$ will still be in effect is given by $\theta^{s-t}$. Therefore, in expectation, we can rewrite the certainty-equivalent Lagrangian replacing $P_{i,s}^c$ with $\theta^{s-t}P_{i,t}^c\left(\Pi_{s-1,t-1}^c\right)^{\gamma_c}\left(\overline{\Pi}_{s,t}\right)^{1-\gamma_c}$, where

$$\Pi_{s,t}^c \equiv \left(\frac{P_{s-1}^c}{P_{t-1}^c}\right)^{\gamma_c}\mathbf{E}_t\left(\prod\nolimits_{j=t}^{s-1}\left(1+\overline{\pi}_j\right)\right)^{1-\gamma_c} \quad s > t.$$

In addition, we can substitute the aggregator's demand function for the output from the $i$th firm:

$$C_{it} = \left(\frac{P_{it}^c}{P_t^c}\right)^{-\epsilon_t^c} C_t,$$

which yields

$$\mathcal{L}_t = \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t}$$

$$\times \left\{ \begin{array}{c} \left(P_{i,t}^c \Pi_{s,t}^c\right)^{1-\epsilon_s^c} (P_s)^{\epsilon_s^c} C_s - W_s H_{is}^c - P_s^{com} COM_{is}^c - P_s^I I_{is}^c - P_s^m M_{is}^c \\ -q_{is}^c \left(K_{i,s+1}^c - (1-\omega) K_{is}^c - I_{is}^c\right) \\ +\lambda_{is} \left( \begin{array}{c} -\left(\frac{P_{i,t}^c \Pi_{s,t}^c}{P_s^c}\right)^{-\epsilon_s^c} C_s + \mathcal{F}\left(L_{is}^c, M_{is}^c, K_{is}^c, u_{is}^c, COM_{is}^c\right) \\ -\frac{\chi_K}{2}\left(\frac{I_{is}^c}{K_{is}^c} - (\omega + g - 1)\right)^2 K_{is}^c \\ -\frac{\chi_I}{2}\left(\frac{I_{is}^c}{I_{i,s-1}^c} - g\right)^2 I_s - \rho\left(1 - e^{\chi_u\left(u_{is}^c - 1\right)}\right)\frac{P_s^I}{P_s^c} K_{is}^c - \frac{\chi_h}{2}\left(\frac{H_s^c}{H_{s-2}^c} - 1\right)^2 A_s \end{array} \right) \end{array} \right\}.$$

To solve for the optimal contract price, set $\frac{\partial \mathcal{L}_t}{\partial P_{i,t}^c} = 0$:

$$\begin{aligned} \frac{\partial \mathcal{L}_t}{\partial P_{i,t}^c} &= \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t}\left(1-\epsilon_s^c\right)\left(\frac{P_{i,t}^c}{P_s}\right)^{-\epsilon_s^c}\left(\Pi_{s,t}^c\right)^{1-\epsilon_s^c} C_s \\ &\quad + \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t}\epsilon_s^c \frac{\lambda_{is}}{P_s^c}\left(\frac{P_{i,t}^c}{P_s^c}\right)^{-\epsilon_s^c - 1}\left(\Pi_{s,t}^c\right)^{-\epsilon_s^c} C_s \\ &= 0 \end{aligned} \tag{A.5}$$

Next, recall that $C_{is} = \left(\frac{P_{it}^c \Pi_{s,t}^c}{P_s^c}\right)^{-\epsilon_s^c} C_s$, so

$$\begin{aligned} & \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t}\left(1-\epsilon_s^c\right)\left(\Pi_{s,t}^c\right) C_{is} + \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t}\epsilon_s^c \frac{\lambda_{is}}{P_{i,t}^c} C_{is} \\ = & \mathbf{E}_t \sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t} C_{is}\left(\left(\epsilon_s^c - 1\right) P_{i,t}^c \Pi_{s,t}^c - \epsilon_s^c \lambda_{is}\right) \\ = & 0, \end{aligned}$$

or, simply,

$$P_{i,t}^c = \mathbf{E}_t \frac{\sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t}\lambda_{is} C_{is}\epsilon_s^c}{\sum_{s=t}^{\infty} \mathcal{R}_{t,s}\theta^{s-t}\left(\Pi_{s,t}^c\right) C_{is}\left(\epsilon_s^c - 1\right)},$$

which is equation (2.24) in the main text.

The Phillips curve in ToTEM is:

$$\widehat{\pi}_t^c = \frac{\gamma}{1+\mathcal{B}\gamma}\left(\widehat{\pi}_{t-1}^c - \Delta \mathbf{E}_t \overline{\pi}_t\right) + \frac{\mathcal{B}}{1+\mathcal{B}\gamma}\mathbf{E}_t \widehat{\pi}_{t+1}^c + \zeta \frac{(1-\theta)(1-\theta\mathcal{B})}{\theta(1+\mathcal{B}\gamma)}\widehat{\lambda}_t^c + \varepsilon_t^p.$$

If we were to assume that capital is costlessly tradable across firms in each period, then $\zeta = 1$. Under our assumption of firm-specific capital with adjustment costs, $\zeta$ will take on a value between 0 and 1. Our procedure for identifying $\zeta$ largely mirrors the derivation in the technical appendix to Altig et al. (2004). The details of our implementation differ slightly because of differences in the production function and the capital utilization costs. The key to identifying $\zeta$ lies in identifying the parameters of the decision rules for the firm-specific relative capital stock and the optimal relative reset price. We posit, and then verify, decision rules of the form:

$$\widehat{k}_{i,t+1}^{c+} = \kappa_1 \widehat{k}_{it}^{c+} + \kappa_2 \widehat{k}_{i,t-1}^{c+} + \kappa_3 \widehat{p}_{it}^c, \tag{A.6}$$

$$\widehat{p}_{it}^{c*} = \widehat{p}_t^{c*} - \psi_0 \widehat{k}_{it}^{c+} - \psi_1 \widehat{k}_{i,t-1}^{c+}, \tag{A.7}$$

where $\widehat{k}_{it}^{c+} = \widehat{k}_{it}^c - \widehat{k}_t^c$ is the relative capital stock of firm $i$, $\widehat{p}_{it}^{c*}$ is firm $i$'s optimal relative reset price, and $\widehat{p}_t^{c*}$ is the average reset price across all resetting firms.

To identify the coefficients in (A.6) and (A.7), we combine these expressions with the firm's first-order conditions for capital, investment, and its reset price. The undetermined-coefficients approach yields five conditions that can be numerically solved for $(\kappa_1, \kappa_2, \kappa_3, \psi_0, \psi_1)$:

$$\begin{aligned}
\widetilde{\gamma}_0\left(\kappa_1 a_2^k + \kappa_2\kappa_1 + \kappa_3 a_2^p\right) + \widetilde{\gamma}_1 a_2^k + \widetilde{\gamma}_2\kappa_1 + \widetilde{\gamma}_3 &= 0, \\
\widetilde{\gamma}_0\left(\kappa_1 a_1^k + \kappa_2^2 + \kappa_3 a_1^p\right) + \widetilde{\gamma}_1 a_1^k + \widetilde{\gamma}_2\kappa_2 + \widetilde{\gamma}_4 &= 0, \\
\Phi\left[\theta - (1-\theta)\psi_0\kappa_3\right] - \widetilde{\gamma}_0\left(\kappa_1 a_0^k + \kappa_2\kappa_3 + \kappa_3 a_0^p\right) - \widetilde{\gamma}_2\kappa_3 - \widetilde{\gamma}_1 a_0^k &= 0,
\end{aligned}$$

$$\begin{aligned}
\psi_0 &= \zeta\phi_1(1-\mathcal{B}\theta)\left[1 + \tau\mathcal{B}\theta A(I-\mathcal{B}\theta A)^{-1}\tau'\right], \\
\psi_1 &= \zeta\phi_1(1-\mathcal{B}\theta)\left[\tau\mathcal{B}\theta A(I-\mathcal{B}\theta A)^{-1}\overline{\tau}'\right],
\end{aligned}$$

where,

$$
\begin{aligned}
a_0^p &\equiv \left\{\theta^2 - 2\theta(1-\theta)\psi_0\kappa_3 - \kappa_3(1-\theta)\left[\psi_0\kappa_1 - (1-\theta)\kappa_3\psi_0^2 + \psi_1\right]\right\}, \\
a_1^p &\equiv -\kappa_2(1-\theta)\left[\theta\psi_0 + \psi_0\kappa_1 - (1-\theta)\kappa_3\psi_0^2 + \psi_1\right], \\
a_2^p &\equiv -(1-\theta)\left\{\begin{array}{c} (\theta\psi_1 + \psi_0\kappa_2 - (1-\theta)\psi_0\kappa_3\psi_1) \\ +\kappa_1\left[\theta\psi_0 + \psi_0\kappa_1 - (1-\theta)\kappa_3\psi_0^2 + \psi_1\right] \end{array}\right\},
\end{aligned}
$$

$$
\begin{aligned}
a_0^k &\equiv \kappa_3\left[\theta + (\kappa_1 - (1-\theta)\kappa_3\psi_0)\right], \\
a_1^k &\equiv \kappa_2\left[\kappa_1 - (1-\theta)\kappa_3\psi_0\right], \\
a_2^k &\equiv \left[(\kappa_2 - (1-\theta)\kappa_3\psi_1) + \kappa_1(\kappa_1 - (1-\theta)\kappa_3\psi_0)\right],
\end{aligned}
$$

$$
\begin{aligned}
\widetilde{\gamma}_0 &\equiv \frac{1-\omega}{1+r}\mathcal{B}\Gamma^{-1}\omega^{-1}, \\
\widetilde{\gamma}_1 &\equiv -\left(\mathcal{B}\Gamma^{-1} + \frac{\chi_K(\omega+g-1)^2}{1+r} + \frac{1-\omega}{1+r}(\Im\Gamma)^{-1} + \frac{1-\omega}{1+r}\mathcal{B}\Gamma^{-1}(1-\omega)\right)\omega^{-1}, \\
\widetilde{\gamma}_2 &\equiv \frac{\chi_K(\omega+g-1)^2}{1+r} + \frac{1-\omega}{1+r}\frac{(1-\Im(1+\mathcal{B}))}{\Im\Gamma} \\
&\quad + \frac{r+\omega}{1+r}\frac{1}{\sigma}\left[1 - b\left(1 + \chi_u\rho\left(\frac{K^c}{C}\right)p^I\right)\right] + \frac{r+\omega}{1+r}\phi_1 \\
&\quad + \left((\Im\Gamma)^{-1} + \frac{1-\omega}{1+r}\Gamma^{-1}\right)\omega^{-1} \\
&\quad + \left(\mathcal{B}\Gamma^{-1} + \frac{\chi_K(\omega+g-1)^2}{1+r} + \frac{1-\omega}{1+r}(\Im\Gamma)^{-1}\right)\left(\frac{1-\omega}{\omega}\right), \\
\widetilde{\gamma}_3 &\equiv -\left(\left((\Im\Gamma)^{-1} + \frac{1-\omega}{1+r}\Gamma^{-1}\right)\omega^{-1}(1-\omega) + \Gamma^{-1}\omega^{-1} + \frac{(1-\Im(1+\mathcal{B}))}{\Im\Gamma}\right) \\
&\quad + (1-\theta)(\psi_0\kappa_1 + \psi_1), \\
\widetilde{\gamma}_4 &\equiv \Gamma^{-1}\omega^{-1}(1-\omega) + (1-\theta)\psi_0\kappa_2, \\
\Phi &\equiv -\left\{\epsilon\frac{r+\omega}{1+r}\frac{1}{\sigma}\left[1 - b\left(1 + \chi_u\rho\left(\frac{K^c}{C}\right)p^I\right)\right] - \frac{r+\omega}{1+r}\phi_2\right\},
\end{aligned}
$$

$$
b \equiv \frac{1}{1 + \rho\sigma + \chi_u\rho\left(\frac{K^c}{C}\right)p^I},
$$

$$\Gamma \equiv \frac{p^I}{\chi_I \lambda g^2}.$$

It can be shown that $\zeta$ is (derivation available from authors):

$$\zeta = \left(1 + \phi_2 + \phi_1 (1 - \mathcal{B}\theta) \left\{ \tau (I - A)^{-1} \left[ \frac{\mathcal{B}\theta}{1 - \mathcal{B}\theta} I - \mathcal{B}\theta A (I - \mathcal{B}\theta A)^{-1} \right] \tau' \right\} \kappa_3 \right)^{-1}, \tag{A.8}$$

where,

$$A \equiv \begin{bmatrix} \kappa_1 & \kappa_2 \\ 1 & 0 \end{bmatrix} : \tau \equiv \begin{bmatrix} 1 & 0 \end{bmatrix},$$

$$\phi_1 = \left[ \frac{\overline{v} - ab}{(1 - \overline{v}) \sigma} \right] : \phi_2 = \epsilon \left[ \frac{\overline{v} - ab}{(1 - \overline{v}) \sigma} \right],$$

$$a = \overline{v} \left[ 1 + \chi_u \rho \left( \frac{K^c}{C} \right) p^I \right] : \overline{v} \equiv \delta_2^{\frac{1}{\sigma}} \left( \frac{K^c}{C} \right)^{\frac{\sigma - 1}{\sigma}} : \delta_2 = (1 - \gamma_c) \upsilon (1 - \delta).$$

## A.2 Households

The instantaneous utility function is given by

$$\mathbb{U}_t = \frac{\mu}{\mu - 1} (C_t^l - \xi C_{t-1}^l)^{\frac{\mu - 1}{\mu}} \exp \left( \frac{\eta (1 - \mu)}{\mu (1 + \eta)} \int_0^1 (E_t N_{ht})^{\frac{1 + \eta}{\eta}} dh \right), \tag{A.9}$$

and the household wishes to maximize $\mathbf{E}_t \sum_{s=t}^{\infty} \boldsymbol{\beta}_{t,s} \mathbb{U}_{hs}^l$ subject to

$$\begin{aligned} & P_{c,t}^{tot} C_t^l + \frac{B_{t+1}}{1 + R_t} + \frac{e_t B_{t+1}^*}{(1 + R_t^*)(1 + \kappa_t)} \\ = \; & B_t + e_t B_t^* + (1 - \varkappa) \left( (1 - \tau_{w,t}) \int_0^1 N_{ht} W_{ht} dh + TF_t \right) + \sum_k \int_0^1 \Delta_{i,k,t} di, \end{aligned} \tag{A.10}$$

The household's problem can be recast in terms of the Lagrangian:

$$\mathcal{L}_t = \mathbf{E}_t \sum_{s=t}^{\infty} \boldsymbol{\beta}_{t,s} \left\{ \begin{array}{c} \frac{\mu}{\mu-1}(C_s^l - \xi C_{s-1}^l)^{\frac{\mu-1}{\mu}} \exp\left(\frac{\eta(1-\mu)}{\mu(1+\eta)}\int_0^1 (E_s N_{hs})^{\frac{1+\eta}{\eta}} dh\right) \\ -\Phi_s \left( \begin{array}{c} P_{c,s}^{tot} C_s^l + \frac{B_{s+1}}{1+R_s} + \frac{e_s B_{s+1}^*}{(1+R_s^*)(1+\kappa_s)} - B_s - e_s B_s^* \\ -(1-\varkappa)\left((1-\tau_{w,s})\int_0^1 N_{hs} W_{hs} dh + TF_s\right) \\ -\sum_k \int_0^1 \Delta_{i,k,s} di \end{array} \right) \end{array} \right\}.$$

Maximizing (2.46) with respect to $C_t^l$ yields:

$$\frac{\partial \mathcal{L}_t}{\partial C_t^l} = \frac{\partial \mathbb{U}_t}{\partial C_t^l} + \beta_{t,t+1}\frac{\partial \mathbb{U}_{t+1}}{\partial C_t^l} - \Phi_t P_{c,t}^{tot} = 0,$$

since, with internal habits, $\frac{\partial \mathbb{U}_{h,t+1}}{\partial C_{ht}^l} \neq 0$. Our particular utility function yields the following:

$$\begin{aligned} P_{c,t}^{tot}\Phi_t &= (C_t^l - \xi C_{t-1}^l)^{-1/\mu} \exp\left(\frac{\eta(1-\mu)}{\mu(1+\eta)}\int_0^1 (E_t N_{ht})^{\frac{1+\eta}{\eta}} dh\right) \\ &\quad -\xi\beta_{t,t+1}(C_{t+1}^l - \xi C_t^l)^{-1/\mu} \exp\left(\frac{\eta(1-\mu)}{\mu(1+\eta)}\int_0^1 (E_{t+1} N_{h,t+1})^{\frac{1+\eta}{\eta}} dh\right) \\ &= \frac{(\mu-1)\,\mathbb{U}_t}{\mu(C_t^l - \xi C_{t-1}^l)} - \xi\beta_{t,t+1}\frac{(\mu-1)\,\mathbb{U}_{t+1}}{\mu(C_{t+1}^l - \xi C_t^l)}. \end{aligned} \tag{A.11}$$

Optimal holdings for foreign and domestic bonds are given by

$$\frac{\partial \mathcal{L}_t}{\partial B_{t+1}} = -\frac{\Phi_t}{1+R_t} + \boldsymbol{\beta}_{t,t+1}\Phi_{t+1} = 0, \tag{A.12}$$

$$\frac{\partial \mathcal{L}}{\partial B_{t+1}^*} = -\frac{\Phi_t e_t}{(1+R_t^*)(1+\kappa_t)} + \boldsymbol{\beta}_{t,t+1}\Phi_{t+1}e_{t+1} = 0, \tag{A.13}$$

which may be combined to yield the uncovered interest parity condition

$$e_t = e_{t+1}\frac{(1+R_t^*)(1+\kappa_t)}{1+R_t}. \tag{A.14}$$

Substituting (A.11) into (A.12) to eliminate $\Phi_t$ and $\Phi_{t+1}$ yields the non-linear form of equation (2.51) in section 2.4.

To compute the optimal wage, we follow the same approach as for prices. First, we substitute in the demand function for labour,

$$N_{hs} = \left(\frac{W_{hs}}{W_s}\right)^{-\epsilon_{w,t}} H_s,$$

and then transform the Lagrangian into its certainty-equivalent counterpart by assigning the probability $(\theta_w)^{s-t}$ that the state $W_{h,s} = \Pi_{w,s,t} W_{h,t}$ will occur in period $s$, $s \geq t$ :

$$\mathcal{L}_t = \sum_{s=t}^{\infty} \beta_{t,s} (\theta_w)^{s-t}$$

$$\times \left\{ \begin{array}{c} \frac{\mu}{\mu-1}(C_s^l - \xi C_{s-1}^l)^{\frac{\mu-1}{\mu}} \\ \times \exp\left( \frac{\eta(1-\mu)}{\mu(1+\eta)} \left( E_s H_s \left(\frac{\Pi_{w,s,t}}{W_s}\right)^{-\epsilon_{w,s}} \right)^{\frac{1+\eta}{\eta}} \int_0^1 (W_{h,t})^{\frac{-\epsilon_{w,s}(1+\eta)}{\eta}} \, dh \right) \\ -\Phi_s \left( \begin{array}{c} P_{c,s}^{tot} C_s^l + \frac{B_{s+1}}{1+R_s} + \frac{e_t B_{s+1}^*}{(1+R_s^*)(1+\kappa)} - B_s - e_s B_s^* \\ -(1-\varkappa)(1-\tau_{w,s}) \Pi_{w,s,t} \left(\frac{W_s}{\Pi_{w,s,t}}\right)^{\epsilon_{w,s}} H_s \int_0^1 (W_{h,t})^{1-\epsilon_{w,s}} \, dh \\ -(1-\varkappa) TF_s - \sum_k \int_0^1 \Delta_{i,k,s} di \end{array} \right) \end{array} \right\}.$$

The first-order condition with respect to $W_{h,t}$, given by $\frac{\partial \mathcal{L}_t}{\partial W_{h,t}} = 0$, is

$$\begin{aligned} 0 \;=\; & \sum_{s=t}^{\infty} \beta_{t,s} (\theta_w)^{s-t} \left\{ \begin{array}{c} \frac{\mu}{\mu-1}(C_s^l - \xi C_{s-1}^l)^{\frac{\mu-1}{\mu}} \exp\left( \frac{\eta(1-\mu)}{\mu(1+\eta)} \int_0^1 (E_s N_{hs})^{\frac{1+\eta}{\eta}} \, dh \right) \\ \times \frac{-\epsilon_{w,s}(1-\mu)}{\mu} \left( E_s H_s \left(\frac{\Pi_{w,s,t}}{W_s}\right)^{-\epsilon_{w,s}} \right)^{\frac{1+\eta}{\eta}} (W_{h,t})^{\frac{-\epsilon_{w,s}(1+\eta)-\eta}{\eta}} \end{array} \right\} \\ & + \sum_{s=t}^{\infty} \beta_{t,s} (\theta_w)^{s-t} \left\{ \begin{array}{c} \Phi_s (1-\varkappa)(1-\tau_{w,s}) \Pi_{w,s,t} \\ \times \left(\frac{W_s}{\Pi_{w,s,t}}\right)^{\epsilon_{w,s}} H_s (1-\epsilon_{w,s}) (W_{h,t})^{-\epsilon_{w,s}} \end{array} \right\}, \end{aligned}$$

or simply

$$
\begin{aligned}
&\sum_{s=t}^{\infty}\beta_{t,s}(\theta_w)^{s-t}\left\{\begin{array}{c}\left(C_s^l-\xi C_{s-1}^l\right)^{\frac{\mu-1}{\mu}}\exp\left(\frac{\eta(1-\mu)}{\mu(1+\eta)}\int_0^1\left(E_sN_{hs}\right)^{\frac{1+\eta}{\eta}}dh\right)\\ \times-\epsilon_{w,s}\left(E_sH_s\left(\frac{\Pi_{w,s,t}}{W_s}\right)^{-\epsilon_{w,s}}\right)^{\frac{1+\eta}{\eta}}\left(W_{h,t}\right)^{\frac{-\epsilon_{w,s}(1+\eta)-\eta}{\eta}}\end{array}\right\}\\
=\;&\sum_{s=t}^{\infty}\beta_{t,s}(\theta_w)^{s-t}\left\{\begin{array}{c}\Phi_s(1-\varkappa)(1-\tau_{w,s})\Pi_{w,s,t}\\ \times\left(\frac{W_s}{\Pi_{w,s,t}}\right)^{\epsilon_{w,s}}H_s(\epsilon_{w,s}-1)\left(W_{h,t}\right)^{-\epsilon_{w,s}}\end{array}\right\}.
\end{aligned}
$$

If we isolate $(W_{h,t})^{-1}$ from the left-hand side, and substitute in the aggregators' demand function for labour, $N_{ht}=\left(\frac{W_{ht}}{W_t}\right)^{-\epsilon_{w,t}}H_t$, we obtain:

$$
W_{h,t}=\mathbf{E}_t\frac{\frac{\mu-1}{\mu}\sum_{s=t}^{\infty}\beta_{t,s}(\theta_w)^{s-t}\mathbb{U}_s\left(E_sN_{hs}\right)^{\frac{1+\eta}{\eta}}\epsilon_{w,s}}{\sum_{s=t}^{\infty}\beta_{t,s}(\theta_w)^{s-t}\Phi_s(1-\varkappa)(1-\tau_{w,s})\Pi_{w,s,t}N_{hs}(\epsilon_{w,s}-1)},
$$

which is equation (2.54) in section 2.4 of the main text.

## A.3 Net foreign-asset holdings

Begin with the budget constraint for lifetime consumers from the main text, given as,

$$
\begin{aligned}
&P_{c,t}^{tot}C_t^l+\frac{B_{t+1}}{1+R_t}+\frac{e_tB_{t+1}^*}{(1+R_t^*)(1+\kappa_t)}\\
=\;&B_t+e_tB_t^*+(1-\varkappa)\left((1-\tau_{w,t})\int_0^1N_{ht}W_{ht}dh+TF_t\right)+\sum_k\int_0^1\Delta_{i,k,t}di,
\end{aligned}
\tag{A.15}
$$

along with

$$
P_{c,t}^{tot}C_t^{tot}=P_{c,t}^{tot}C_t^c+P_{c,t}^{tot}C_t^l,\tag{A.16}
$$

$$
P_{c,t}^{tot}C_t^c=\varkappa\left((1-\tau_{w,t})W_tN_t+TF_t\right),\tag{A.17}
$$

$$W_tH_t = (1-\varkappa)\int_0^1 N_{ht}W_{ht}dh + \varkappa W_tN_t, \tag{A.18}$$

$$B_{t+1} = (1+R_t)\left\{B_t + \left(P_t^g G_t + TF_t - TX_t\right)\right\}, \tag{A.19}$$

$$TX_t = \tau_{w,t}W_tH_t + \tau_{c,t}\left(\frac{P_{c,t}^{tot}C_t^{tot}}{1+\tau_{c,t}}\right). \tag{A.20}$$

Substitute equations (A.16), (A.17) and (A.18) into (A.15), which gives

$$\begin{aligned} & P_{c,t}^{tot}C_t^{tot} + \frac{B_{t+1}}{1+R_t} + \frac{e_tB_{t+1}^*}{(1+R_t^*)(1+\kappa_t)} \\ = \quad & B_{h,t} + e_tB_{h,t}^* + (1-\tau_{w,t})W_tH_t + TF_t + \sum_k \int_0^1 \Delta_{i,k,t}di. \end{aligned} \tag{A.21}$$

This yields the budget constraint for consumption by both types of household. Next, substitute equation (A.19) into the above equation, which gives,

$$\begin{aligned} & P_{c,t}^{tot}C_t^{tot} + \frac{e_tB_{t+1}^*}{(1+R_t^*)(1+\kappa_t)} \\ = \quad & e_tB_{h,t}^* + (1-\tau_{w,t})W_tH_t - P_t^gG_t + TX_t + \sum_k \int_0^1 \Delta_{i,k,t}di. \end{aligned}$$

Then substitute equation (A.20) to eliminate taxes from the budget constraint:

$$\begin{aligned} & \frac{e_tB_{t+1}^*}{(1+R_t^*)(1+\kappa_t)} \\ = \quad & e_tB_{h,t}^* + W_tH_t - P_t^gG_t - \frac{P_{c,t}^{tot}c_t^{tot}}{1+\tau_{c,t}} + \sum_k \int_0^1 \Delta_{i,k,t}di. \end{aligned} \tag{A.22}$$

Finally, we must substitute the definition of dividends, for all sectors, into the above equation. Dividends are simply the accounting profits for the $ith$ firm for each sector. Using core consumption as an example, we have

$$\Delta_{i,c,t} = P_{it}^cC_{it} - W_tH_{it}^c - P_t^{com}COM_{it}^c - P_t^II_{it}^c - P_t^mM_{it}^c, \tag{A.23}$$

so total profits in the core consumption sector are

$$\begin{aligned}\int_0^1 \Delta_{i,c,t} di &= \int_0^1 P_{it}^c C_{it} di - W_t \int_0^1 H_{it}^c di - P_t^{com} \int_0^1 COM_{it}^c di \\ &\quad - P_t^I \int_0^1 I_{it}^c di - P_t^m \int_0^1 M_{it}^c di.\end{aligned}$$

Profits in the investment, government, and non-commodity export sector are defined similarly. Profits for the commodity producer are given as

$$\Delta_{\cdot,com,t} = COM_t e_t P_{com,t}^* - W_s H_s^{com} - P_s^I I_s^{com}.$$

Profits for the import/commodity distribution sector are

$$\begin{aligned}\Delta_{\cdot,dist,t} &= \left(P_t^m - e_t P_t^*\right) M_t \\ &\quad + \left(P_t^{com} - e_t P_{com,t}^*\right)\left(C_t^{com} + COM_t^c + COM_t^I + COM_t^g + COM_t^{nx}\right).\end{aligned} \tag{A.24}$$

Note also that total nominal consumption can be expressed as

$$P_{c,t}^{tot} C_t^{tot} = \left(1 + \tau_{c,t}\right)\left(P_t^c C_t + P_t^{com} C_t^{com}\right), \tag{A.25}$$

and

$$COM_t = C_t^{com} + COM_t^c + COM_t^I + COM_t^g + COM_t^{ncx} + X_{com,t}. \tag{A.26}$$

If we combine the profit conditions for the consumption, investment, government, commodity, and finished-product export, along with the following market-clearing conditions,

$$\begin{aligned}W_t H_t &= W_t \left(H_t^c + H_t^I + H_t^g + H_t^{ncx} + H_t^{com}\right), \\ P_t^I I_t &= P_t^I \left(I_t^c + I_t^I + I_t^g + I_t^{ncx} + I_t^{com}\right), \\ P_t^m M_t &= P_t^m \left(M_t^c + M_t^I + M_t^g + M_t^{ncx}\right),\end{aligned}$$

we obtain

$$\begin{aligned}\sum_k \int_0^1 \Delta_{i,k,t} di &= P_t^c C_t + P_t^g G_t + P_t^x X_{nc,t} + COM_t e_t P_{com,t}^* \\ &\quad - W_t H_t + P_t^{com} C_t^{com} - e_t P_t^* M_t \\ &\quad - e_t P_{com,t}^* \left( C_t^{com} + COM_t^c + COM_t^I + COM_t^g + COM_t^{ncx} \right).\end{aligned}$$

Next, substitute equations (A.25) and (A.26) to yield

$$\sum_k \int_0^1 \Delta_{i,k,t} di = \frac{P_{c,t}^{tot} C_t^{tot}}{1+\tau_{c,t}} + P_t^g G_t + P_t^x X_{nc,t} + e_t P_{com,t}^* X_{com,t} - W_t H_t - e_t P_t^* M_t. \quad \text{(A.27)}$$

Finally, substitute (A.27) into (A.22) to yield

$$\frac{e_t B_{t+1}^*}{\left(1+R_t^*\right)\left(1+\kappa_t\right)} = e_t B_{h,t}^* + P_t^x X_{nc,t} + e_t P_{com,t}^* X_{com,t} - e_t P_t^* M_t. \quad \text{(A.28)}$$

Equation (A.28) can then be expressed in Canadian-dollar-denominated net foreign assets, multiplying through by the correctly dated value of the exchange rate:

$$e_{t+1} B_{t+1}^* = \frac{e_{t+1}}{e_t} \left(1+R_t^*\right)\left(1+\kappa_t\right) \left\{ e_t B_{h,t}^* + P_t^x X_{nc,t} + e_t P_{com,t}^* X_{com,t} - e_t P_t^* M_t \right\}, \quad \text{(A.29)}$$

which is equation (2.78) in the main text.

# Appendix B

# The Output Gap in ToTEM

In ToTEM, there is no production function for GDP per se; rather, there are production functions for the components of the national accounts identity:

$$Y_t = C_t + I_t + G_t + X_{com,t} + X_{nc,t} - M_t. \tag{B.1}$$

In addition, all imports are treated as inputs to the production of the consumption, investment, government, and non-commodity export goods, and there is also a role for non-exported commodities in production. These inputs are combined in three distinct stages of production, to allow for different elasticities of substitution across the four different production inputs:

$$\mathcal{G}(A_t L_{it}, u_{it} K_{it}) = \left( (\alpha_{c,1})^{\frac{1}{\sigma}} (A_t L_{it})^{\frac{\sigma-1}{\sigma}} + (1-\alpha_{c,1})^{\frac{1}{\sigma}} (u_{it} K_{it})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{B.2}$$

$$\mathcal{H}(\mathcal{G}, COM_{it}^c) = ((\alpha_{c,2})^{\frac{1}{\varrho}} \mathcal{G}(A_t L_{it}^c, u_{it}^c K_{it}^c)^{\frac{\varrho-1}{\varrho}} + (1-\alpha_{c,2})^{\frac{1}{\varrho}} (COM_{it}^c)^{\frac{\varrho-1}{\varrho}})^{\frac{\varrho}{\varrho-1}}, \tag{B.3}$$

$$\mathcal{F}(\mathcal{H}, M_{it}^c) = \left( (\alpha_{c,3})^{\frac{1}{\varphi}} (\mathcal{H}(\mathcal{G}, COM_{it}^c))^{\frac{\varphi-1}{\varphi}} + (1-\alpha_{c,3})^{\frac{1}{\varphi}} (M_{it}^c)^{\frac{\varphi-1}{\varphi}} \right)^{\frac{\varphi}{\varphi-1}}, \tag{B.4}$$

to use the consumption-goods sector as an example. Stage-one production (equation (B.2)) may be thought of as pure value added, where the second combines value added with commodities and the third combines that composite with imports. Once linearized around steady state, equations (B.2), (B.3), and (B.4) may be combined to yield, using the consumption sector as an example,

$$\ln C_t \approx \phi_1(\ln A_t + \ln L_t^c) + \phi_2(\ln K_t^c + u_t^c) + \phi_3 \ln COM_t^c + \phi_4 \ln M_t^c,$$

where $\sum \phi_i = 1$, and $\phi_i$ depends non-linearly on the steady-state nominal share of the $i$th input in total production, and the elasticities of substitution.

The linearized production functions can then be combined with the linearized market-clearing conditions to yield an approximate expression for output, given by

$$\ln Y_t \approx \varkappa_1 \ln A_t + \varkappa_2 \ln A_t^i + \sum_{j=1}^{5} \zeta_j \ln L_{j,t} + \sum_{j=1}^{5} \rho_j \left(\ln K_{j,t} + u_{j,t}\right), \tag{B.5}$$

where $\ln A_t$ ($\ln A_t^i$) is the log of economy-wide (investment-sector-specific[1]) labour-augmenting technology (LAT),[2] and $u_{j,t}$ is capital utilization. Potential output is then defined as[3]

$$\ln \widetilde{Y}_t \equiv \varkappa_1 \ln A_t + \varkappa_2 \ln A_t^i + \sum_{j=1}^{5} \zeta_j \ln \widetilde{H}_{j,t} + \sum_{j=1}^{5} \rho_j \ln K_{j,t}. \tag{B.6}$$

Then the output gap is simply a composite of the total labour input (effort and hours) and capital utilization gaps in each production sector:

$$\widehat{Y}_t \equiv \ln Y_t - \ln \widetilde{Y}_t = \sum_{j=1}^{5} \zeta_j \widehat{L}_{j,t} + \sum_{j=1}^{5} \rho_j \widehat{u}_{j,t}. \tag{B.7}$$

[1] Currently in ToTEM, sector-specific technology is employed only in the investment sector of the model.

[2] Recall that technology must be labour augmenting only with CES technology, to ensure a stationary capital-to-output ratio in the presence of non-stationary technology.

[3] The steady-state or potential levels for effort and capital utilization in each sector are normalized to one.

# Bank of Canada Technical Reports
## *Rapports techniques de la Banque du Canada*

**Technical reports are generally published in the language of the author, with an abstract in both official languages and are available on the bank's website (see bottom of page).** ***Les rapports techniques sont généralement publiés dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue. On peut les consulter dans le site Web de la Banque du Canada, dont l'adresse est indiquée au bas de la page.***



Copies and a complete list of working papers are available from:
*Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à* :


Publications Distribution, Bank of Canada
234 Wellington Street, Ottawa, Ontario K1A 0G9
Telephone: 1 877 782-8248
(toll free in North America)
Email: publications@bankofcanada.ca
Website: http://www.bankofcanada.ca

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9
Téléphone : 1 877 782-8248 (sans frais en Amérique du Nord)
Adresse électronique : publications@banqueducanada.ca
Site Web : http://www.banqueducanada.ca