

信息摩擦、信号处理与货币政策

邓燕飞* 沈吉† 张军‡

内容提要：粘价格和粘性信息两种理论孰优孰劣是宏观经济学中颇有争论的话题，从拟合内生惯性的方面来看粘性价格模型更好，从模拟外生惯性的角度而言粘性信息模型更优。当粘性概率与非粘性概率平方之比等于噪音方差与冲击方差之比时，粘价格可作为理性疏忽的特例。本文证明，当更新信息的概率等于卡尔曼增益时，粘性信息亦可作为理性疏忽的特例。用中国数据拟合理性疏忽菲利普斯曲线后，得到了较粘性价格模型更好的拟合效果；在标准的三方程动态理性疏忽一般均衡框架下，也得到了类似于粘性信息模型生成的通货膨胀等宏观经济变量的符合数据特征的脉冲响应图。据此，兼具粘价格和粘性信息各自优点的理性疏忽可以替换这两者，采用货币政策分析框架时也可避免在这两者之间犹疑不决。就实际可操作的货币政策层面而言，理性疏忽框架中最优简单规则下的福利损失更小，更靠近理想的最优货币政策。

关键词：理性疏忽；粘性信息；粘价格；最优货币政策；最优简单规则

一 引言

习近平总书记在关于《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》的说明中强调指出，“考虑到未来一个时期外部环境不稳定不确定因素较多，存在不少可能冲击国内经济发展的风险隐患，新冠肺炎疫情全球大流行影响深远，世界经济可能持续低迷，中长期规划目标要更加注重经济结构优化，引导各方面把工作重点放在提高发展质量和效益上”（习近平，2020）。具体到货币政策的出台与实施上，无疑也会是一个巨大挑战。新凯恩斯理论框架是各国央行基于“货币非中性”的前提制订有效货币政策的有力工具，但对价格变动机理未作更深入探究的粘价格（sticky prices）以及巧妙避开信号提取或信息处理的粘性信息（sticky information）在日趋复杂的国内外形势下用于货币政策分析会力不从心，有必要转向能专门处理不稳定、不确定因素的理论——理性疏忽（rational inattention）。本文将从模拟外生惯性、拟合内生惯性、测算福利损失等三个角度论证理性疏忽都比粘性信息和粘价格更适用于分析理论上的最优货币政策及实际可操作层面的最优简单规则。

就粘价格和粘性信息这一组理论而言，文献中一直存在后者能否替代前者的讨论。¹各自来看，在拟合内生惯性方面粘价格模型更好，而在模拟外生惯性方面，粘性信息模型更优。有较好应用价值

*男，1983-，讲师（浙江财经大学经济学院），电子邮箱：dengyf@zufe.edu.cn。本文入选第二届中国宏观经济学者论坛，薛涧坡教授等学者提供了建设性意见和建议；研究过程中还得到 Kristoffer Nimark, Hassan Afrouzi, Choongryul Yang, M. Murat Aslan、Luigi Paciello, Mirko Wiederholt, Bartosz Mackowiak, Johanner Pfeifer 等人的热情帮助，一并致以诚挚谢意。

†男，1982-，副教授（北京大学光华管理学院），通讯作者，电子邮箱：jishen@gsm.pku.edu.cn。

‡男，1963-，文科资深教授、博导、院长（复旦大学经济学院，中国经济研究中心），电子邮箱：junzh_2000@fudan.edu.cn。

¹Mankiw and Reis (2002) 提出粘性信息理论可以替代粘价格理论已近廿载，国外对这支文献有较大力度的关注，后续学者们或以其他方式进一步验证或通过其他视角否定或在原有理论基础上改进的研究成果已有不少，国内亦有学者做了一些研究（王军，2009；彭兴韵，2011；卞志村等，2016；邓燕飞，2018）。从有限的范围了解的情况看，目前国内基于新凯恩斯 DSGE 框架的各种应用性研究，多未加讨论地假设经济主体按确定性交错（Taylor, 1979）或随机交错方式（Calvo, 1983）调整价格或工资（即粘价格或粘性工资等名义刚性理论），继而在此逻辑中推演

(比如货币政策分析)的理论模型应兼备这两方面的优势,否则,“错误的模型将导致错误的结论”(Ball et al., 2005; Kitamura, 2008; Angeletos et al., 2020)。Dupor et al. (2010) 遂将粘性价格和粘性信息合成为双粘性模型,呈现内、外生惯性的效果确实更好。美中不足的是,双粘性模型并非通过微观经济主体的行为优化及均衡条件严谨推导而来,微观基础不牢。²

Woodford (2002); Paciello et al. (2014) 研究表明,当粘性概率与非粘性概率平方之比等于噪音方差与冲击方差之比时,粘性价格可作为理性疏忽(rational inattention)这一信息摩擦类模型的特例。若能探明粘性信息亦可或在何条件下可纳入理性疏忽的框架,则理性疏忽模型有望像双粘性模型那样兼具粘性价格和粘性信息分别在拟合内生惯性和模拟外生惯性方面的优点,但又不用像双粘性模型那样生嵌硬套,而可通过将信息理论应用到成熟的 DSGE 框架中严格推导建立(Sims, 2010; Maćkowiak et al., 2009, 2015; Maćkowiak et al., 2018a; Afrouzi et al., 2020),这将有助应用该类模型在愈加纷繁复杂的背景下权衡信息的成本收益,增强政策研究结论的可信度。中国人民银行行长易纲在 2020 金融街论坛上的讲话指出,2020 年以来,新冠疫情对全球经济带来严重冲击,中国货币政策响应将及时有力、灵活有度(易纲, 2020),而此目标的实现也将要求对经济主体决策所依赖的信息及其信息处理方式、信息处理能力等进行深度分析。

经过理论推导和两相对比,本文探明,当更新信息的概率等于卡尔曼增益时,粘性信息理论的确可作为理性疏忽模型的特例。用中国数据拟合理性疏忽菲利普斯曲线时得到了类似于粘性价格模型优于粘性信息模型的拟合效果,且在一般均衡框架中模拟成本加成冲击、需求冲击及货币政策冲击后得到了类似于粘性信息模型产生的符合数据特征的脉冲响应。更重要地,基于最优货币政策和最优简单规则的分析框架所测算的基准福利损失和相对福利损失值从定量的角度表明理性疏忽较粘性信息模型更适用于货币政策分析。

粘性价格假设决策信息完美,粘性信息假设决策信息以一定概率完美。³理性疏忽直面信息不完美,且较信号提取还能权衡信息处理的成本收益以使经济主体的注意力实现最优配置。从面临不完全信息时的信号提取,到完全信息理性预期的粘性价格,到每期以一定概率获得完全信息的粘性信息,再到处理不完全信息的理性疏忽,信息摩擦理论经历了波澜起伏的发展阶段:

(1) 不完全信息与理性人假设

Friedman and Schwartz (1963) 从美国近一百年的样本数据中研究认为通货膨胀与失业之间存在实质为“货币非中性”的替代关系,但限于短期。随后, Friedman (1968) 对短期菲利普斯曲线的形成作出理论解释,他认为:当货币供给突然增加时,价格水平上升,实际货币工资下降,劳动力成本降低,企业愿意雇佣更多员工,而员工只关注名义工资因而错误地认为工资报酬上涨,其工作意愿加强,产出增加,即短期出现“货币非中性”与员工基于不完全信息作出决策密切相关。Lucas (1972, 1973) 假设劳

得到相应结论。纵然对粘性信息也仍有异议(如 Coibion, 2010),若不加讨论地在应用性研究中直接采用名义刚性机制,很可能得到有失偏颇的研究结果。Mankiw and Reis (2002) 的分析着重于模拟名义冲击后通货膨胀、产出缺口等宏观经济变量的动态路径,而 Arslan (2008) 认为模拟效果取决于设定理想的参数值,校准的参数自然也可通过贝叶斯估计获得,但如此得来的参数值却又不一定能使粘性信息模型更好地呈现通货膨胀等宏观变量符合普遍看法的外生惯性,因此单从模拟外生惯性的角度来谈论模型好坏也确有值得商榷和完善之处,更何况粘性信息菲利普斯曲线拟合内生惯性的效果还要另加审视。换言之,从已有研究成果来看,粘性价格和粘性信息各有利弊,很难判断这两个理论模型哪个更好。为此, Dupor et al. (2010) 假设有些企业不会调整价格而有些企业会根据当前或过去的信息调整价格,而作出价格调整的企业一定比例根据新近信息,相应比例根据过去信息,如此嵌套而成了一个双粘性通货膨胀方程。从模拟外生惯性和拟合内生惯性的两方面来看,都较粘性价格和粘性信息模型有更好的表现。但这并非基于部门优化及均衡条件严格推导而来,因此限制了双粘性模型作出可靠的货币政策分析的能力。自然,更为理想的模型应是既对内、外生惯性有较好的解释力,又有牢靠的微观基础以具备提高预测准确性的潜力。另可参看: Keen (2007); Kiley (2007); Klenow et al. (2007); Korenok (2008); Trabandt (2009); Knotek II (2010); Arslan (2010); Eggertsson et al. (2019)。

²此等不足类似于 Gali et al. (1999); Christiano et al. (2005) 两个模型中凭空而生的滞后通货膨胀一项。

³在微观博弈论中,完全信息(full information)与完美信息(perfect information)是有区别的概念,而在宏观理论中,完全信息与完美信息皆指对决策信息的全部掌握,是语义相同的概念。

资双方都是理性人构建的不完全信息迭代模型及在自耕农经济体中建立的“信号提取”模型，对货币中性等问题做了更丰富的讨论。

着眼于不完全信息状态下的信号提取模型在活跃了一段时间后，有所沉寂，原因之一是噪音的扰动外生给定，并且经济主体处理信息的成本收益未加考量，换言之，信息处理过程并不严谨 (Sims, 1998)。

(2) 基于完全信息的理性预期

信号提取模型是严谨处理信息不完美的一次重要尝试，但上述不足促使完全信息理性预期模型兴起。加速膨胀的菲利普斯曲线出现后⁴，吸收理性预期假设，相继发展了粘性价格菲利普斯曲线、粘性信息菲利普斯曲线，双粘性通货膨胀方程等新凯恩斯理论模型。⁵

基于Taylor (1980); Rotemberg (1982); Calvo (1983) 等研究成果, Rotemberg et al. (1997); Clarida et al. (1999) 等学者发展了名义刚性下的传统粘性价格理论，完善为新凯恩斯菲利普斯曲线，弥补了传统模型缺乏微观基础的不足。⁶但人们逐渐意识到它的三个弊端：一，预先公布的可信任的抑制通货膨胀政策会导致价格继续上升这一奇怪结果；二，它不能很好地解释通货膨胀惯性；三，它不能很好地解释为什么货币政策冲击对通货膨胀具有滞后和逐步的影响。

Mankiw and Reis (2002) 进一步从传统粘性价格理论的微观基础出发，基于经济主体获取信息、分析信息等过程中因成本产生而使宏观经济环境的信息在大众中逐步缓慢传播的假设，提出用粘性信息菲利普斯曲线（特点是滞后理性预期）替代新凯恩斯菲利普斯曲线（特点是前瞻理性预期）以更好地刻画通货膨胀动态。⁷通过模拟，他们发现粘性信息菲利普斯曲线可以弥合上述不足。虽然粘性信息理论在模拟通货膨胀对货币政策冲击的反应上有良好表现，但Coibion (2010) 从拟合实际通货膨胀数据的角度，指出粘性价格模型更好。

为弥补粘性价格模型基于完全信息理性预期假设的不足，另一种尝试是考虑部分企业是适应性预期，即这部分企业设定其产品价格的依据是上一期平均重修价格外加滞后通货膨胀率 (Gali and Gertler, 1999; 范从来 and 高洁超, 2016)，或部分企业根据滞后通货膨胀率改变价格 (Christiano et al., 2005)，这些被称为混合新凯恩斯菲利普斯曲线（特点是包含了理性预期和非理性预期），它能较好地拟合通货膨胀数据。Dupor et al. (2010) 则提出了可以比肩混合新凯恩斯菲利普斯曲线的双粘性通货膨胀方程（特点是兼具不同形式的理性预期），用美国数据的匹配结果也很理想。

Ball et al. (2005) 指出，既然以适应性预期为特征的加速膨胀的菲利普斯曲线和以理性预期为特征的新凯恩斯菲利普斯曲线都存有缺点，将这两者融为一体的混合新凯恩斯菲利普斯曲线很有可能兼具两者之不足，并提出，由于加速膨胀的菲利普斯曲线本质上是非理性预期在主导人们的预期行为，即使数据拟合得更好，但由于非理性本身的“随意性”，会导致由此分析的政策结论难有较高的可信度。从这个角度却也突显了双粘性通货膨胀方程的优势：不含非理性预期，但数据拟合效果与包含非理性预期的混合新凯恩斯菲利普斯曲线相当 (Dupor et al., 2010)。

遗憾的是，简单机械地将粘性价格和粘性信息嵌套而成的双粘性模型虽然兼具粘性价格和粘性信息

⁴加速膨胀的菲利普斯曲线又被称为附加预期的菲利普斯曲线，认为经济主体会吸收以往经验用于未来预期。该曲线产生的货币政策效应符合人们的惯常看法，也与实证结果较为吻合 (Gordon, 1997; Staiger et al., 1997)；但可能陷入卢卡斯批判，即当政策规则调整时，外生的非理性的适应性预期不能调整以对应新的制度 (Ball et al., 2005)。

⁵粘性价格和粘性信息同为前瞻预期，但后者还兼有滞后预期的特点，双粘性则是将粘性价格和粘性信息这两种理性预期嵌为一体，如此构造了异质性预期；还有一种方法是将非理性预期与理性预期糅为异质性预期。参看李拉亚 (1993); 肖争艳 等 (2005); 许志伟 等 (2015); 邓燕飞 等 (2017)。

⁶陈彦斌 (2008); 齐鹰飞 (2011); 卞志村 等 (2016); 姜峰 (2016) 等学者研究了中国菲利普斯曲线。

⁷另可参看国内学者王军 (2009); 彭兴韵 (2011) 的综述。

这两个理论的长处，但由于它并非基于部门最优和均衡条件严谨推导而来，微观基础并不明朗。

(3) 信息不完全的理性疏忽

粘价格等理性预期模型中的预期方式不管如何演变，始终局限于“完全信息的条件预期”。粘性信息理论虽也立足于完全信息的理性预期，但又对偏离完全信息有所考量，它假设每期以 κ 概率获得完全信息，相应地， $1 - \kappa$ 概率未能获得。如此，既引入了信息摩擦，又巧妙地避开了信号提取或信息处理等相对复杂的技术问题。

理性疏忽重新转到“不完全信息的条件预期”上。所谓条件，即经济主体决策时对状态变量掌握的信息。不同于“信号提取”假设噪音的波动外生，理性疏忽进一步将波动内生，借鉴信息理论中的熵、互信息等技术手段可在 DSGE 框架中权衡信息处理的成本收益，在此框架内可研究消费、投资、经济周期等诸多议题。⁸

Sims (2010); Luo et al. (2013); MaćKowiak et al. (2018b) 对理性疏忽学说的形成与发展有全面介绍。众所周之，经济学的核心议题是如何分配稀缺资源，通常又指有形资源，而理性疏忽转而聚集研究如何分配注意力这一无形的稀缺资源 (Wiederholt, 2010)。以企业部门为例，假设不存在战略互补性这样的反馈机制，完全信息状态下企业的理想价格与名义总需求水涨船高。完全信息指企业对名义总需求历史及现今信息的完全掌握。但放松这一假设，意味着企业定价将基于一个包含噪音的状态变量（名义总需求）的序列。当企业投入更多注意力于状态变量以掌握更准确的数据时，企业的定价将更靠近理想价格，但由于企业的注意力是稀缺资源，在投入注意力掌握更多信息以降低对状态变量的不确定性时，也增加了企业的成本。若假设新息、噪音都服从正态分布，则容易在线性二次高斯框架下，求解企业注意力的最优配置 (Peng et al., 2006; Maćkowiak et al., 2009; Paciello, 2012; Melosi, 2014; Pasten et al., 2016; Matějka, 2016; Afrouzi, 2020)。若假设家庭部门是理性疏忽的，则可求解在收入之用于消费或投资上注意力的最优分配 (Sims, 2003; Luo, 2008; Luo et al., 2017; Tutino, 2013; Zorn, 2016)。

Paciello and Wiederholt (2014) 分别讨论了内、外生信息背景下对于成本加成冲击和技术冲击的最优货币政策。Maćkowiak and Wiederholt (2015) 建立了理性疏忽为特点的动态随机一般均衡框架，发现该模型对于货币政策冲击及技术冲击的脉冲响应与 VAR 实证脉冲响应匹配得很好。Mondria et al. (2010); Luo et al. (2012) 将理性疏忽约束应用于开放经济领域。Hébert et al. (2018); Miao (2019) 有别于多数研究的离散时间处理方式，转向连续时间。Miao et al. (2019); Afrouzi et al. (2020) 在更一般化的多维世界中对理性疏忽的求解提供了新思路。Yang (2019) 在企业有多个产品的决策环境下构建了理性疏忽与菜单成本的混合模型。国内学者对理性疏忽有过综述性质的介绍 (王军, 2011)，对于该理论在经济周期中的作用及与货币政策的关系也有过探究 (于泽, 2009; 李拉亚, 2011; 王军等, 2013)。陈雨露 and 马勇 (2013) 指出，“当现实世界已经出现了新的变化，而在旧的事态空间下发展起来的理论又不足以预见和反映这种变化时，理论的革新就成了必然选择。”陈雨露 (2015) 进一步认为，2008 年金融危机之后，尝试在“金融—实体经济”内生性框架下重建宏观经济学理论的努力逐渐成为一种潮流和趋势。理性疏忽的宏观理论将信息结构内生是搭建“金融—实体经济”内生性框架的重要一环。

(4) 为何重回信息不完全

理论模型为何推陈出新？无外乎是为更好地解释历史现实及较好地预测未来。基于此，文献中常见的对新凯恩斯宏观理论模型的评判依据是：先是拟合内生惯性的效果。用基于某理论推导而来的菲利普斯曲线，拟合通货膨胀的真实数据，若能更好拟合，则该理论更佳 (Kiley, 2007; Coibion, 2010)。这是由于数据中的通货膨胀本身存在惯性，因此一个理论生成的通货膨胀方程应有呈现通货膨胀惯性的机制，

⁸除了理性疏忽之外，也有从测度等角度讨论对理性预期的偏离情况 (参看 钟春平等, 2015; 何启志等, 2017)。

比如附加预期的菲利普斯曲线通过引入适应性预期根据拇指规则硬生地添加了一个通货膨胀滞后项，而粘性价格菲利普斯曲线与粘性信息菲利普斯曲线则是通过引入不同形式的理性预期内生出了通货膨胀滞后项；再是模拟外生惯性的功力。考虑不同冲击时由菲利普斯曲线与需求曲线构成的完整经济系统在均衡状态下通货膨胀等宏观变量的脉冲响应，比如“名义冲击”（主要指货币政策冲击）下的通货膨胀反应会有所延迟且呈驼峰状 (Mankiw et al., 2002; Christiano et al., 2010)，而“非名义冲击”（比如技术冲击、成本加成冲击）下的通货膨胀反应更迅速 (Dupor et al., 2009; Coibion, 2010; Carrillo, 2012)。

以此为评判依据，上述完全信息理性预期模型多在某一方面或该方面的某点上有较好表现，而在另一方面或其他点上表现欠佳 (Ball et al., 2005)。央行若欲基于理论模型通过承诺或相机抉择控制利率或货币供给以平滑冲击，前提是模型应能正确反应宏观经济变量内在的运动状态，并能呈现外生冲击后的应有路径。那是否存在同时符合上述两方面判断标准的建模方法？完全信息与现实脱离过远，信号提取对信息不完全的处理过于简单，信息处理受限且能权衡信息处理成本收益的理性疏忽模型有望同时具备上述两方面优势。⁹

以上研究鲜有推理性疏忽菲利普斯曲线从内生惯性的角度验证拟合效果，也没有研究将理性疏忽与粘性信息这一组信息摩擦在统一框架内从外生惯性的角度通过脉冲响应做比对研究，完成以上两方面的工作有助建立理性疏忽在最优货币政策分析上的优势，这将为央行决策提供更完善的理论依据，亦可有助完善和丰富新凯恩斯理论文献。为此，下一节将在找到粘性信息菲利普斯曲线与动态理性疏忽菲利普斯曲线的内在联系后，用中国数据做实证检验。第三节建立动态随机一般均衡框架模拟成本加成冲击、需求冲击及货币政策冲击的脉冲响应。第四节基于 Ramsey 问题和最优简单规则通过测算相对福利损失值对理性疏忽和粘性信息进一步比较。最后是小结。

二 理性疏忽菲利普斯曲线的实证检验

理性疏忽与粘性信息的思想形成于上世纪末本世纪初，Mankiw and Reis (2002)、Woodford (2002)、Sims (2003) 等理论创始人或主要贡献者就这两者的异同都有扼要一提，简言之，就决策所依赖的信息而言，粘性信息理论假设每期有一定的概率可以得到最新信息，而理性疏忽学说认为每期的信息中都存在噪音。

如表1所示，在一个相似的框架中，从信息完美到信息不完美（从右向左）的演化过程一目了然。简介如下：

需求侧：由假设货币流通速度标准化为 1 的货币数量方程表示，其中表示货币总供给或名义 GDP 的 m_t 为外生变量，服从 AR(1) 过程，外生冲击项 ϵ_t 是均值为 0 方差为 σ_ϵ^2 的白噪声。

供给侧：起点都是合意定价方程 $\hat{p}_{it}^* = \hat{p}_t + \alpha_y \tilde{y}_t$ ，即第 i 家企业在完全信息状态下等于边际成本的价格。其中， \hat{p}_t 表示总价格水平， $\tilde{y}_t = \hat{y}_t - \hat{y}_t^f$ 表示产出缺口，简便起见，本节假设自然产出率 \hat{y}_t^f 为 0，因此 $\tilde{y}_t = \hat{y}_t$ ， $\alpha_y \in (0, 1]$ 越小意味着越强的定价战略互补性或真实刚性。¹⁰

信息完美或信息摩擦之下，并未假设价格刚性，因此总价格水平为 $\hat{p}_t = \int_0^1 \hat{p}_{it} di$ ， \hat{p}_{it} 为第 i 家企业设定的价格水平。信息完美时自然有 $\hat{p}_{it} = \hat{p}_{it}^*$ 。而当存在信息摩擦时， $\hat{p}_{it} \neq \hat{p}_{it}^*$ ：（1）粘性信息是种假设存在信息摩擦又巧妙避开了信号提取或信息处理的一种设定，因此用 $\hat{p}_{it,h} = \mathbb{E}_{t-h} \hat{p}_{it}^*$ 表示基于 h 期前的

⁹对于 Dong and Wen (2019), Dong et al. (2020), 董丰 and 许志伟 (2020), 费兆奇 and 刘康 (2020) 中讨论的资产泡沫、金融系统性风险、国债市场风险定价等问题也可从理性疏忽的角度加以考量。

¹⁰全文用小写字母表示内、外生变量的对数值，小写字母上加“~”表示相应内生变量偏离稳态的百分比，假设所有外生变量没有趋势增长，因此稳态处外生变量的对数值为 0。

二 理性疏忽菲利普斯曲线的实证检验

完美信息设定价格；(2) 如前所述，理性疏忽直面信息不完美，直接刻画存在噪音成分的观测变量，即 $s_{it} = m_t + \xi_{it}$ ，其中 ξ_{it} 是均值为 0 方差为 σ_ξ^2 的白噪声。¹¹ 早期的信号提取模型通常假设 σ_ξ^2 外生，理性疏忽假设其为内生，根据净利润损失最小化可求解最优注意力。

表 1: 不完全信息理性预期与完全信息理性预期系统^①

理性疏忽 噪音的波动内生	噪音信息 噪音的波动外生	粘性信息 一定概率无噪音	完全信息 无噪音
$\min_{\sigma_\xi^2} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\pi_{11}}{2} (\hat{p}_{it}^* - \hat{p}_{it})^2 + f(\mu) \right],$ ^①	$\hat{p}_t = \int_0^1 \hat{p}_{it} di,$ ^②	$\hat{p}_t = \sum_{h=0}^{\infty} \hat{p}_{it,h},$	$\hat{p}_t = \int_0^1 \hat{p}_{it} di,$
$\hat{p}_{it}^* = \hat{p}_t + \alpha_y \tilde{y}_t,$ $\hat{p}_{it} = \mathbb{E}[\hat{p}_{it}^* I_i^t],$	$\hat{p}_{it}^* = \hat{p}_t + \alpha_y \tilde{y}_t,$ $\hat{p}_{it} = \mathbb{E}[\hat{p}_{it}^* I_i^t],$	$\hat{p}_{it}^* = \hat{p}_t + \alpha_y \tilde{y}_t,$ $\hat{p}_{it,h} = \mathbb{E}_{t-h} \hat{p}_{it}^*,$	$\hat{p}_{it}^* = \hat{p}_t + \alpha_y \tilde{y}_t,$ $\hat{p}_{it} = \hat{p}_{it}^*,$
$s_{it} = m_t + \xi_{it},$	$s_{it} = m_t + \xi_{it},$	$s_{it} = m_t,$	$s_{it} = m_t,$
$m_t = \hat{p}_t + \tilde{y}_t,$	$m_t = \hat{p}_t + \tilde{y}_t,$	$m_t = \hat{p}_t + \tilde{y}_t,$	$m_t = \hat{p}_t + \tilde{y}_t,$
$\Delta m_t = \Delta m_{t-1} + \epsilon_t,$	$\Delta m_t = \Delta m_{t-1} + \epsilon_t,$	$\Delta m_t = \Delta m_{t-1} + \epsilon_t.$	$\Delta m_t = \Delta m_{t-1} + \epsilon_t.$
$\log \sigma_{m t-1}^2 - \log \sigma_{m t}^2 \leq \mu.$	$\xi_{it} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\xi^2).$ ^②	$\epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2).$ ^③	$I_i^t \equiv I_i^{-1} \cup \{s_{i0}, s_{i1}, \dots, s_{it}\}.$ ^④

① 本表的方程板块纵向看共四列：第一列是具备内生信息结构的理性疏忽假设，即观测方程中的噪音 ξ_{it} 其二阶矩内生；第二列是外生信息结构的噪音信息学说，前述二阶矩外生；第三列是以某个概率获得完全信息的粘性信息理论；第四列是基于完全信息的理性预期模型。本表的方程板块横向看共七行：第一行从左到右第一个方程是理性疏忽系统中定价主体的行为方程，经济含义为投入更多注意力将使噪音减少而使定价 \hat{p}_{it} 尽可能靠近利润最大化时的合意定价 \hat{p}_{it}^* ，但投入的注意力有一定成本，成本函数为 $f(\mu)$ ，第一行从左到右第二、三、四个方程分别是理性疏忽（同于噪音信息）、粘性信息、完全信息时的总价格水平 \hat{p}_t 的定义；第二行四个方程相同，皆为利润最大化时的合意定价方程；第三行从左到右第一、二个方程相同，为理性疏忽系统中的定价行为，即基于截止到第 t 时期的信息集 I_i^t 的定价，第三行第三个方程是粘性信息理论中假设每期以 $1 - \kappa$ 的概率获得完全信息时的定价方程，第三行第四个方程表示完全信息时的定价亦即为合意定价；第四行从左到右第一、二个方程相同，为理性疏忽系统中的信噪方程，货币 m_t 是状态变量，观测误差为 ξ_{it} ，第四行第三、四个方程相同，皆为对 m_t 的完全掌握；第五行相同，为总需求曲线；第六行相同，为货币供给增长率服从 AR(1) 的状态转移方程；第七行第一个方程为理性疏忽系统中内生信息结构下存在约束的互信息，第七行第二、三个式子分别是假设观测误差及模型误差是白噪声，第七行第四个式子是对信息集的定义。

② 该方程借鉴 [Paciello and Wiederholt \(2014\)](#) 脚注 3，可假设一个线性的注意力成本函数，令 $f(\mu) = c\mu$ ，常数 $c > 0$ ；

③ 该式第 1、2 列共用；

④ 该式第 1、2、3 列共用；

⑤ 该式第 1、2、3、4 列共用。

粘性信息菲利普斯曲线的拟合效果不佳 ([Coibion, 2010](#))。本节将首先简要证明粘性信息菲利普斯曲线可作为理性疏忽菲利普斯曲线的特例及其成立的条件。继而将进一步拓展更一般意义的动态理性疏忽菲利普斯曲线，并用中国数据进行实证拟合，以从内生惯性的角度显示理性疏忽的优势。

(1) 粘性信息菲利普斯曲线

[Mankiw and Reis \(2002\)](#) 假设每家企业每期以 κ 的概率获得完全信息，由此巧妙地避开了存在信息摩擦时需进行信号提取等复杂问题，则第 t 期企业定价为

$$\hat{p}_{it,h} = \kappa(1 - \kappa)^h \mathbb{E}_{t-h} \hat{p}_{it}^*.$$

总价格水平 $\hat{p}_t \equiv \int_0^1 \hat{p}_{it,h} di$ ，可进一步表示为

$$\hat{p}_t = \kappa \sum_{h=0}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} \hat{p}_t^*.$$

根据 [Mankiw and Reis \(2002\)](#) 的附录或 [Menz and Vogel \(2009\)](#)，不难得到

$$\hat{\pi}_t = \kappa \sum_{h=0}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-1-h} (\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) + \frac{\kappa}{1 - \kappa} \alpha_y \tilde{y}_t,$$

¹¹ 外生变量 m_t 服从 AR(1) 过程，提取最优信号的观测方程才有如此形式。(详见 [MaćKowiak et al., 2016](#))

其中, $\Delta \tilde{y}_t = \tilde{y}_t - \tilde{y}_{t-1}$ 是产出增长率。

如果在第 $t-1$ 时随机抽取一个样本询问其关于第 t 时的状态变量的预期, 平均而言, 应为:

$$\bar{\mathbb{E}}_{t-1}(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) \equiv \kappa \sum_{h=0}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-1-h}(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t).$$

因此得到以下形式的粘性信息菲利普斯曲线:

$$\hat{\pi}_t = \bar{\mathbb{E}}_{t-1}(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) + \frac{\kappa}{1 - \kappa} \alpha_y \tilde{y}_t.$$

笔者此前已有专文对粘性信息理论进行实证检验, 相应内容不再作为本文重点。但与接下来的理性疏忽作一比较, 故此, 参数估计与拟合结果一并列于后文表2。为方便计, 上述方程稍加调整为

$$\hat{\pi}_t = \bar{\mathbb{E}}_{t-1} \hat{\pi}_t + b_1 \bar{\mathbb{E}}_{t-1} \Delta \tilde{y}_t + b_2 \tilde{y}_t, \quad (2.1)$$

其中, 待估参数 $b_1 \equiv \alpha_y$, $b_2 \equiv \frac{\kappa}{1 - \kappa} \alpha_y$ 。

(2) 噪音信息菲利普斯曲线

理性疏忽学说与早期Lucas (1972, 1973) 的信号提取模型一脉相承, 不同之一在于, 早期为通过贝叶斯法则对后验分布进行更新完善的静态问题, 后逐步发展为通过卡尔曼滤波更新信息的动态问题; 无论静态还是动态, 此前噪音的波动都被假设为常数, 后考察内生即关注最优注意力分配的问题。

对于企业 i , 其定价为

$$\hat{p}_{it} = \mathbb{E}(\hat{p}_{it}^* | I_i^t),$$

其中信息集为 $I_i^t = I_i^{t-1} \cup \{s_{i0}, s_{i1}, \dots, s_{i,t-1}, s_{it}\}$ 。

借鉴Woodford (2002); MaćKowiak et al. (2016), 不妨猜想 $\hat{p}_t = H\hat{p}_{t-1} + Gm_{t-1} + B\epsilon_t$, 并令 \mathbf{e}_i 为第 i 个元素为 1 其余元素为 0 的向量, 将其与货币供给的运动规则与信噪方程一起写成状态空间表达式:

$$\begin{aligned} \mathbf{x}_t &= \mathcal{F}\mathbf{x}_{t-1} + \mathcal{A}\epsilon_t, \\ s_i^t &= \mathbf{e}_i' \mathbf{x}_t + \xi_{it}, \\ \mathbf{x}_t &\equiv \begin{bmatrix} m_t \\ p_t \end{bmatrix}, \quad \mathcal{F} \equiv \begin{bmatrix} 1 & 0 \\ G & H \end{bmatrix}, \quad \mathcal{A} \equiv \begin{bmatrix} 1 \\ B \end{bmatrix}. \end{aligned}$$

将需求方程代入合意定价方程, 单个企业的价格调整方程变为:

$$\hat{p}_{it} = \mathbb{E}[\alpha_y m_t + (1 - \alpha_y) \hat{p}_t | I_i^t].$$

总价格水平因此为:

$$\hat{p}_t = \alpha_y m_{t|t} + (1 - \alpha_y) \hat{p}_{t|t},$$

其中下标 “ $t|t$ ” 表示基于加总的直到第 t 期的个体信息对第 t 期相应变量的平均估计。

根据 Kalman 滤波可以得到企业 i 对当前状态向量的最优估计¹²

$$\mathbf{x}_{t|t}(i) = \mathbf{x}_{t|t-1}(i) + \kappa [s_i^t - \mathbf{e}_i' \mathbf{x}_{t|t-1}],$$

¹²Hamilton (1994)pp.372-

其中, κ 是卡尔曼增益矩阵。如前所述, 初始状态向量由无条件期望产生, 即 $\mathbf{x}_{t|t-1}(i) = \mathcal{F}\mathbf{x}_{t-1|t-1}(i)$, 将状态方程和观测方程一起代入上式并加总 (注意到 $\int_0^1 \xi_{it} di = 0$)

$$\begin{aligned}\mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + \kappa \mathbf{e}'_1 (\mathbf{x}_t - \mathbf{x}_{t|t-1}), \\ &= \mathcal{F}\mathbf{x}_{t-1|t-1} + \kappa \mathbf{e}'_1 [(\mathcal{F}\mathbf{x}_{t-1} + \mathcal{A}\epsilon_t) - \mathcal{F}\mathbf{x}_{t-1|t-1}], \\ &= \kappa \mathbf{e}'_1 \mathcal{F}\mathbf{x}_{t-1} + (\mathbf{I} - \kappa \mathbf{e}'_1) \mathcal{F}\mathbf{x}_{t-1|t-1} + \kappa \mathbf{e}'_1 \mathcal{A}\epsilon_t.\end{aligned}$$

将总价格水平改写为:

$$\begin{aligned}\hat{p}_t &= \begin{bmatrix} \alpha_y & (1 - \alpha_y) \end{bmatrix} \begin{bmatrix} m_{t|t} \\ \hat{p}_{t|t} \end{bmatrix} \equiv \bar{\phi} \mathbf{x}_{t|t}, \\ &= \bar{\phi} [\kappa \mathbf{e}'_1 \mathcal{F}\mathbf{x}_{t-1} + (\mathbf{I} - \kappa \mathbf{e}'_1) \mathcal{F}\mathbf{x}_{t-1|t-1} + \kappa \mathbf{e}'_1 \mathcal{A}\epsilon_t], \\ &= \kappa \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \rho & 0 \\ G & H \end{bmatrix} \begin{bmatrix} m_{t-1} \\ \hat{p}_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_y & (1 - \alpha_y) \end{bmatrix} \begin{bmatrix} \rho & 0 \\ G & H \end{bmatrix} \begin{bmatrix} m_{t-1|t-1} \\ \hat{p}_{t-1|t-1} \end{bmatrix} \\ &\quad - \kappa \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \rho & 0 \\ G & H \end{bmatrix} \begin{bmatrix} m_{t-1|t-1} \\ \hat{p}_{t-1|t-1} \end{bmatrix} + \kappa \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ B \end{bmatrix} \epsilon_t, \\ &= Gm_{t-1} + H\hat{p}_{t-1} + B\epsilon_t.\end{aligned}$$

其中, $\kappa \equiv \bar{\phi} \kappa$ (注意, 前者 κ 未加粗表示标量, 后者 κ 加粗表示向量)。对比上述最后两式, 待定系数定为:

$$\begin{cases} G = \kappa, \\ H = 1 - \kappa, \\ B = \kappa. \end{cases}$$

显见:

$$\hat{p}_t = (1 - \kappa)\hat{p}_{t-1} + \kappa m_{t-1} + \kappa \epsilon_t.$$

现在要确定卡尔曼增益 κ 。

$$\begin{aligned}\kappa &\equiv \mathbf{MSE}_{t|t-1}^x \mathbf{e}_1 (\mathbf{e}'_1 \mathbf{MSE}_{t|t-1}^x \mathbf{e}_1 + \sigma_\xi^2)^{-1} = \Sigma \mathbf{e}_1 (\mathbf{e}'_1 \Sigma \mathbf{e}_1 + \sigma_\xi^2)^{-1}, \\ &= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sigma_\xi^2 \right)^{-1} = \begin{bmatrix} \frac{\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2} \\ \frac{\Sigma_{21}}{\Sigma_{11} + \sigma_\xi^2} \end{bmatrix},\end{aligned}$$

其中, $\mathbf{MSE}_{t|t-1}^x$ 是基于截至第 $t-1$ 期的信息集对第 t 期的 x 作预期的均方误, 设为矩阵 Σ , 各元素分别为 $\Sigma_{11}, \Sigma_{12}, \Sigma_{21}, \Sigma_{22}$ 。向量 κ 的两个元素自然为: $\kappa_1 = \frac{\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2}$, $\kappa_2 = \frac{\Sigma_{21}}{\Sigma_{11} + \sigma_\xi^2}$ 。

因此问题转向到求解 Σ_{11} 和 Σ_{21} 。参看附录B, 可知 (简化符号起见, 令 $q \equiv \frac{\sigma_\epsilon^2}{\sigma_\xi^2}$, $X \equiv \sqrt{1 + 4q}$):

$$\begin{cases} \Sigma_{11} = \sigma_\epsilon^2 \frac{X+1}{2}; \\ \Sigma_{21} = \sigma_\epsilon^2 \frac{(X+1)+2q}{(X-1)+\frac{2}{\kappa}}. \end{cases}$$

据此可得 κ_1 和 κ_2 。又因 $\kappa = \alpha_y \kappa_1 + (1 - \alpha_y) \kappa_2$ ，所以 $q\kappa^2 + \alpha_y \kappa - \alpha_y = 0$ 。 κ 有两个根，根据需求函数及状态方程， $\tilde{y}_t = (1 - \kappa)\tilde{y}_{t-1} + (1 - \kappa)\epsilon_t$ ，显见，收敛而非发散的 \tilde{y}_t 要求 $\kappa \in (0, 1]$ 。解得（令 $\hat{q} \equiv \frac{\alpha_y}{q} = \alpha_y \frac{\sigma_\epsilon^2}{\sigma_\xi^2} > 0$ ）：

$$\kappa = \frac{\sqrt{\hat{q}^2 + 4\hat{q}} - \hat{q}}{2}.$$

重回定价方程：

$$\hat{p}_{it} = \mathbb{E}(\hat{p}_{it}^* | I_i^{t-1}) + \kappa[s_{it} - \mathbb{E}(s_{it} | I_i^{t-1})].$$

假设所有企业的卡尔曼增益相同，则根据总价格水平的定义有：

$$\begin{aligned} \hat{p}_t &= \int_0^1 \hat{p}_{it} di = \int_0^1 \mathbb{E}(\hat{p}_{it}^* | I_i^t) di, \\ &= \int_0^1 \mathbb{E}(\hat{p}_{it}^* | I_i^{t-1}) di + \kappa \int_0^1 [s_{it} - \mathbb{E}(s_{it} | I_i^{t-1})] di, \\ &= \int_0^1 \mathbb{E}(\hat{p}_{it}^* | I_i^{t-1}) di + \kappa[\hat{p}_{it}^* - \mathbb{E}(\hat{p}_{it}^* | I_i^{t-1})], \end{aligned}$$

上述最后一步推导用了白噪声的性质，即 $\int_0^1 \epsilon_{it} di = 0$ 。此外，观测方程中的扰动项 ϵ_{it} 与观测变量正交，因此 $\int_0^1 s_{it} di = \hat{p}_{it}^*$ ，并且 $\mathbb{E}(s_{it} | I_i^{t-1}) = \mathbb{E}(\hat{p}_{it}^* | I_i^{t-1})$ 。

进一步，两边同时减去 $\kappa \hat{p}_t$ ，并稍作变换，可得：

$$\hat{p}_t = \int_0^1 \mathbb{E}(\hat{p}_{it}^* | I_i^{t-1}) di + \frac{\kappa}{1 - \kappa}(\hat{p}_{it}^* - \hat{p}_t).$$

注意到 $\hat{p}_{it} = \mathbb{E}(\hat{p}_{it}^* | I_i^t)$ ， $\hat{p}_{t-1} = \int_0^1 \hat{p}_{i,t-1} di$ ，因此 $\hat{p}_{t-1} = \int_0^1 \mathbb{E}(\hat{p}_{i,t-1}^* | I_i^{t-1}) di$ 。于上式两边同减该式，并根据已知条件 $\hat{\pi}_t = \hat{p}_t - \hat{p}_{t-1}$ 、 $\hat{p}_{it}^* = \hat{p}_t + \alpha_y \tilde{y}_t$ 和 $\hat{p}_{it}^* - \hat{p}_{i,t-1}^* = \hat{\pi}_t + \alpha_y \Delta \tilde{y}_t$ ，稍作运算可得

$$\hat{\pi}_t = \int_0^1 \mathbb{E}[(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) | I_i^{t-1}] di + \frac{\kappa}{1 - \kappa} \alpha_y \tilde{y}_t.$$

相似地，可以定义第 $t-1$ 期对第 t 期状态变量的平均预期为：

$$\bar{\mathbb{E}}_{t-1}(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) \equiv \int_0^1 \mathbb{E}[(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) | I_i^{t-1}] di.$$

最后得到形式上完全同于粘性信息菲利普斯曲线的理性疏忽菲利普斯曲线，

$$\hat{\pi}_t = \bar{\mathbb{E}}_{t-1}(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) + \frac{\kappa}{1 - \kappa} \alpha_y \tilde{y}_t.$$

显见，不同之处在于 κ 在前者中表示更新信息的概率，而后者表示卡尔曼增益。¹³

上述由更新信息的概率变为卡尔曼增益后，也即由粘性信息菲利普斯曲线转变成了噪音菲利普斯曲线，亦即噪音的扰动外生条件下的理性疏忽菲利普斯曲线。此时，投入注意力减少噪音波动的成本未加

¹³此亦为 Afrouzi et al. (2018) 的一个结论，然而该文献只有一个技术性概要。在 Afrouzi 的提示帮助下，本文对此结论的完整过程予以呈现。

考量。更一般化地, 假设理性疏忽中噪音的波动内生, Afrouzi and Yang (2018) 建立了动态理性疏忽菲利普斯曲线:

$$\hat{\pi}_t = \bar{\mathbb{E}}_{t-1}(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t) + \delta_0 \alpha_y \tilde{y}_t - \sum_{\tau=1}^{\infty} \beta^\tau \delta_\tau [\bar{\mathbb{E}}_t(\hat{\pi}_{t+\tau} + \alpha_y \Delta \tilde{y}_{t+\tau}) - \mathbb{E}_t(\hat{\pi}_{t+\tau} + \alpha_y \Delta \tilde{y}_{t+\tau})].$$

最后综括号中那项表示相较于完全信息的预测误差。当 $\beta = 0$ 时, 动态理性疏忽菲利普斯曲线退化为噪音菲利普斯曲线。有必要指出的是, 上述方程形式上与 Dupor et al. (2010) 建立的双粘性方程比较相像, 都包含了滞后预期和前瞻预期, 但前瞻项并非是预测误差, 只是完全信息下的预期变量。

Coibion et al. (2018a)^{p.1466, 表 5} 提供了上述方程的一个稍有不同的简洁版本以方便估计和检验, 稍作调整为

$$\hat{\pi}_t = \bar{\mathbb{E}}_{t-1} \hat{\pi}_t + b_1 \bar{\mathbb{E}}_{t-1} \Delta \tilde{y}_t + b_2 \tilde{y}_t + b_3 [\bar{\mathbb{E}}_t(\hat{\pi}_{t+1} + \Delta \tilde{y}_{t+1}) - \hat{i}_t]. \quad (2.2)$$

(3) 检验理性疏忽菲利普斯曲线

方程 (2.2) 中待估参数 $b = \{b_1, b_2, b_3\}$ 是一种简化形式, 并非结构性的, 理论上的定性判断是: $b_1 > 0$, $b_2 > 0$, $b_3 < 0$ 。方程中涉及的变量有三类: 可观测 (通货膨胀、产出水平、名义利率)、可推测 (潜在产出) 和可调查 (通货膨胀预期、产出增长率预期)。理性疏忽学说的关键特征是并非基于完全信息的理性预期假说, 所以检验这一理论不宜通过 VAR 估计完全信息理性预期值。由于预期调查的受访者难以掌握完全信息, 因而检验理性疏忽菲利普斯曲线将选用相应变量预期的调查数据。¹⁴

除了通货膨胀预期的调查数据, 还需产出增长率预期的调查数据。前者而言, 张成思 and 党超 (2016) 于附录提供了 2000 年第一季度至 2014 年第三季度基于中国人民银行调查数据测算的同比预期通胀率; 后者来说, 用中国人民银行城镇储户问卷调查指数表中就业预期指数 (季度数据) 作为代理变量。¹⁵

为匹配上述较难得到的预期调查序列数据, 同样时间段的产出水平和名义利率从 Chang et al. (2015) 进行了细致研究和重新测算 (均已经过季节性调整) 的中国宏观数据库中提取。得到产出缺口依赖于对潜在产出的估计, 可选择二次去势或 HP 滤波。如若将驱动变量选为实际边际成本, 则可考虑用劳动收入份额作为代理变量, 亦可从该数据库下载使用。

Dupor et al. (2010) 使用的两步法的估计思路于此可借鉴。不同之处在于, 第一步直接将表示预期调查的相关变量 (上标 “s”) 代入理性疏忽菲利普斯曲线, 得到通货膨胀理论值 $\hat{\pi}_t^m(b)$:

$$\hat{\pi}_t^m(b) = \hat{\pi}_{t-1}^s + b_1 \Delta \tilde{y}_{t-1}^s + b_2 \tilde{y}_t + b_3 [(\hat{\pi}_{t+1}^s + \Delta \tilde{y}_{t+1}^s) - \hat{i}_t].$$

第二步是与实际通货膨胀进行比较, 选择最小化两者距离的方差 (“var”) 的参数, 即

$$\hat{b} = \arg \min_b \text{var}(\hat{\pi}_t - \hat{\pi}_t^m).$$

粘性信息理论假设以 κ 的概率获得完全信息, 因此同样用上述两步法对粘性信息菲利普斯曲线的估计在执行程序上稍有不同, 集中在第一步上, 转而通过向量自回归 (VAR) 表示相关变量完全信息下的前瞻及滞后理性预期 (参看 Dupor et al., 2010; Woodford, 2001; Mavroeidis et al., 2014)。

¹⁴ 不同形式的预期尤其完全信息理性预期和非完全信息调查预期对于的实证表现可参看 Nunes (2010), Fuhrer (2012), Mavroeidis et al. (2014), Coibion and Gorodnichenko (2012), Coibion et al. (2015), Coibion et al. (2018a), Coibion et al. (2018b), Coibion et al. (2020)。

¹⁵ 数据来源: <http://www.pbc.gov.cn/diaochatongjisi/116219/116227/index.html>

估计结果如表2所示：首先，定性上看，已估参数的符号同于理论判断， b_1 和 b_2 皆大于 0 而 b_3 小于 0。由于待估方程未含常数项，所以非中心调整的 R^2 表示理性疏忽菲利普斯曲线有显著更佳的拟合效果，达到 0.963，通货膨胀的理论值与实际值距离的方差 var-e 也小得多，仅为 0.086；而粘性信息菲利普斯曲线的拟合优度和方差分别为 0.431 和 1.306。就本节从拟合内生惯性的角度比较粘性信息菲利普斯曲线与理性疏忽菲利普斯曲线的目标而言，结果显示，理性疏忽菲利普斯曲线显著更优，图1对这一结果提供了更为直观的证据。

表 2: 理性疏忽 (RI) 和粘性信息 (SI) 模型的参数估计和拟合结果

	b_1	b_2	b_3	R^2	var-e
RI	0.356** [0.0934,0.2042]	1.374** [1.5167,10.9161]	-0.361*** [-0.2600,-0.4754]	0.963	0.086
SI	0.18** [-0.0632,0.4157]	4.56** [-0.0277,11.0645]	-	0.431	1.306

¹ 样本期间为 2000Q1 至 2014 年 Q2；“***”表示在 0.01 水平上显著，“**”表示在 0.05 水平上显著，“*”表示在 0.10 水平上显著。

² 第一、二行分别为理性疏忽菲利普斯曲线和粘性信息菲利普斯曲线的估计结果。第 2、3、4 列分别为待估参数的点估计及区间估计值；第五列是拟合优度 (R^2)；第六列是模型通货膨胀预测序列与实际通货膨胀序列间差距的标准差 (var-e)。

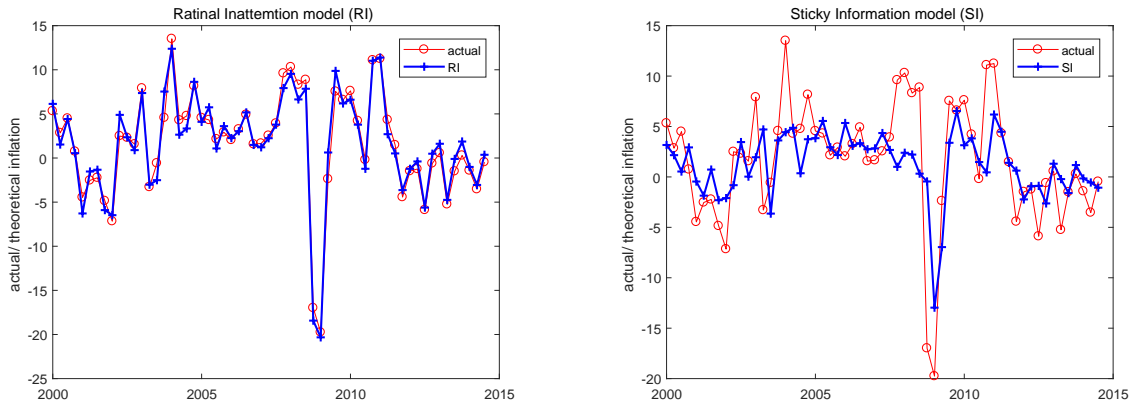


图 1: 理性疏忽 (RI) 和粘性信息 (SI) 的拟合效果。理论值分别基于方程2.2(Rational Inattention, RI) 和2.1(Sticky Information, SI)，实际值 (actual) 源自张成思 and 党超 (2016) 于附录提供的 2000 年第一季度至 2014 年第三季度基于中国人民银行调查数据测算的同比预期通胀率。

此处并没做稳健性检验，但Dupor et al. (2010) 一文得出的将粘性价格和粘性信息嵌套为双粘性模型的实证检验结果显著优于单纯的粘性信息模型的研究结论可作为侧面呼应，因为其构建的双粘性模型与动态理性疏忽模型形态上相似，都包含了前瞻预期和滞后预期，所不同之处在于：(1) 双粘性模型中的前瞻、滞后预期皆为基于完全信息的理性预期，而动态理性疏忽模型中的前瞻、滞后预期皆为非完全信息的理性疏忽；(2) 双粘性模型为简单机械性质的嵌套构成，微观基础欠缺，而动态理性疏忽模型根据标准化的 DSGE 理论的构建步骤而成，具备牢固的微观基础。虽有这些差异，但结构上的相似性决定两者在比较粘性信息模型的优势上可一定程度的相互验证。

三 理性疏忽模型的均衡动态

若假设家庭部门拥有完美信息，则动态理性疏忽一般均衡模型容易构建。虽然上节已推导出更新信息的概率为卡尔曼增益时，粘性信息菲利普斯曲线与理性疏忽菲利普斯曲线一致，亦即为 $\beta = 0$ 条件下的动态理性疏忽菲利普斯曲线。但粘性信息与理性疏忽的内部机制仍有很大区别，因此本节将考察理性疏忽模型能否与粘性信息模型一样生成“驼峰”状明显的通货膨胀惯性。Afrouzi and Yang (2020) 建立了一个简单的三方程动态理性疏忽系统，以研究技术冲击和货币政策冲击之下产出和通货膨胀等宏观经济变量的脉冲响应，本节稍作变化，转向分析成本加成冲击、总需求冲击与货币政策冲击下通货膨胀、产出缺口等关键宏观经济变量的均衡动态。沿用调整过的三方程如下：

表 3: 动态随机一般均衡系统

理性疏忽	粘性信息
$\hat{p}_{it} = \mathbb{E}_{it}[(\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t) I_t^i],$	$\hat{p}_t = \kappa \sum_{h=0}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h}(\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t),$
$\tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma} (\tilde{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}) + u_t,$	$\tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma} (\tilde{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}) + u_t,$
$\tilde{i}_t^e = \rho \tilde{i}_{t-1}^e + (1 - \rho)(\phi_\pi \hat{\pi}_t + \phi_y \tilde{y}_t^e) + v_t.$	$\tilde{i}_t^e = \rho \tilde{i}_{t-1}^e + (1 - \rho)(\phi_\pi \hat{\pi}_t + \phi_y \tilde{y}_t^e) + v_t.$

¹ 本表纵向看共两列：第一列是加入利率规则得以构成一个闭合的理性疏忽经济系统；第二列是增加利率规则后构成一个闭合的粘性信息经济系统。本表的方程板块横向看共三行：第一行从左到右分别是相应理论中供给侧的行为方程，第二行是需求侧的行为方程，第三行是相同的利率规则。

注意到由于两部门信息设定不同，因此预期符号有差异， \mathbb{E}^\times 和 \mathbb{E} 分别表示基于不完全信息与完全信息的预期算子。从附录 A 可知，此时 $\tilde{y}_t^e \equiv \hat{y}_t - \hat{y}_t^e$, $\hat{y}_t^e = \frac{\alpha_a}{\alpha_y} a_t$, $\tilde{i}_t^e \equiv \hat{i}_t - \hat{i}_t^e$, $\hat{i}_t^e = \frac{\sigma \alpha_a}{\alpha_y} \mathbb{E}_t \Delta a_{t+1}$ 。外生变量 z_t 、 u_t 、 v_t 皆为 AR(1) 过程： $z_t = \rho_z z_{t-1} + \epsilon_t^z$, $u_t = \rho_u u_{t-1} + \epsilon_t^u$, $v_t = \rho_v v_{t-1} + \epsilon_t^v$ 。其中扰动项 $\epsilon_t^z, \epsilon_t^u, \epsilon_t^v$ 都是均值为 0、方差分别为 $\sigma_z^2, \sigma_u^2, \sigma_v^2$ 的白噪声。令 $\epsilon_t \equiv (\epsilon_t^z, \epsilon_t^u, \epsilon_t^v)'$, $\tilde{\epsilon}_t \equiv (\epsilon_t^z, \epsilon_t^u, \tilde{\epsilon}_t^v)'$, $\mathbf{x}_t \equiv (\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-L+1})'$, $\tilde{\mathbf{x}}_t \equiv (\tilde{\epsilon}_t, \tilde{\epsilon}_{t-1}, \dots, \tilde{\epsilon}_{t-L+1})'$, 则

$$\tilde{\mathbf{x}}_t = (\mathbf{I} - \mathbf{\Lambda} \mathbf{M}')^{-1} \mathbf{x}_t,$$

其中 \mathbf{I} 是 3×3 的单位阵， $\mathbf{\Lambda}$ 是主对角线上 $(3i, 3i)$ 为 1 其余皆为 0 的对角阵， $i = 1, 2, \dots, L$, L 表示截断的滞后阶数， \mathbf{M} 是一个下位移矩阵， $\mathbf{M}_{ij} = \delta_{i, j+3}$ (Kronecker delta 函数)。¹⁶

对三个冲击的 AR(1) 过程迭代后可知：

$$\begin{cases} z_t = \mathbf{H}'_z \mathbf{x}_t, & \mathbf{H}'_z = (1, 0, 0, \rho_z, 0, 0, \rho_z^2, 0, 0, \dots, \rho_z^{L-1}, 0, 0); \\ u_t = \mathbf{H}'_u \mathbf{x}_t, & \mathbf{H}'_u = (0, 1, 0, 0, \rho_u, 0, 0, \rho_u^2, 0, \dots, 0, \rho_u^{L-1}, 0); \\ v_t = \mathbf{H}'_v \mathbf{x}_t, & \mathbf{H}'_v = (0, 0, 1, 0, 0, \rho_v, 0, 0, \rho_v^2, \dots, 0, 0, \rho_v^{L-1}). \end{cases}$$

猜想无摩擦时能使利润最大化的理想价格的表达式为：

$$\begin{aligned} \hat{p}_{it}^* &= \phi_z z_t + \phi_u u_t + \phi_v v_t, \\ &= \Phi_z(L) \epsilon_t^z + \Phi_u(L) \epsilon_t^u + \Phi_v(L) \epsilon_t^v, \\ &= (\phi_z \mathbf{H}'_z + \phi_u \mathbf{H}'_u + \phi_v \mathbf{H}'_v) \mathbf{x}_t, \\ &= (\phi_z \mathbf{H}'_z + \phi_u \mathbf{H}'_u + \phi_v \mathbf{H}'_v) (\mathbf{I} - \mathbf{\Lambda} \mathbf{M}') \tilde{\mathbf{x}}_t, \end{aligned}$$

¹⁶ https://en.wikipedia.org/wiki/Shift_matrix

其中 ϕ_z, ϕ_u, ϕ_v 皆为待定系数, $\Phi_z(L), \Phi_u(L), \Phi_v(L)$ 都是滞后多项式。采用 Afrouzi and Yang (2020) 提出的求解思路¹⁷, 注意到货币政策冲击引起价格变动的部分是非平稳的, 因此提取单位根并稍作整理:

$$\begin{aligned}\hat{p}_{it}^* &= \Phi_z(L)\epsilon_t^z + \Phi_u(L)\epsilon_t^u + (1-L)\Phi_v(L)(1-L)^{-1}\epsilon_t^v, \\ &= \Phi_z(L)\epsilon_t^z + \Phi_u(L)\epsilon_t^u + \tilde{\Phi}_v(L)\tilde{\epsilon}_t^v, \\ &= \mathbf{H}'\tilde{\mathbf{x}}_t,\end{aligned}$$

其中, $\tilde{\Phi}_v(L) \equiv (1-L)\Phi_v(L)$, 而 $\tilde{\epsilon}^v \equiv (1-L)^{-1}\epsilon_t^v = (1+L+L^2+\cdots+L^\infty)\epsilon_t^v = \sum_{i=0}^{\infty} \epsilon_{t-i}^v$; ¹⁸ $\mathbf{H}'_z, \mathbf{H}'_u, \mathbf{H}'_v$ 已知, \mathbf{H}' 待定, 迭代逼近求解该内生变量时可设定初始值 $\mathbf{H}'_0 = \mathbf{H}'_z + \mathbf{H}'_u + \mathbf{H}'_v$ 。

综上, 有如下状态空间表达式:

$$\begin{cases} \tilde{\mathbf{x}}_t = \mathbf{A}\tilde{\mathbf{x}}_{t-1} + \mathbf{Q}\epsilon_t, \\ \hat{p}_{it}^* = \mathbf{H}'\tilde{\mathbf{x}}_t, \end{cases}$$

其中,

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

注意到只有价格是非平稳序列, 通货膨胀、产出缺口和利率水平都是平稳序列。又因 $\hat{p}_t = \int_0^1 \hat{p}_{it} di$, $\hat{p}_{it} = \mathbb{E}_{it}^\times \hat{p}_{it}^*$, 所以

$$\begin{aligned}\hat{p}_t &= \mathbf{H}' \int_0^1 \mathbb{E}_{it}^\times \mathbf{x}_t di \equiv \mathbf{H}'_p \tilde{\mathbf{x}}_t, \\ \hat{\pi}_t &= \mathbf{H}'_p (\mathbf{I} - \mathbf{A}\mathbf{M}')^{-1} (\mathbf{I} - \mathbf{M}') \mathbf{x}_t \equiv \mathbf{H}'_\pi \mathbf{x}_t,\end{aligned}$$

其中 $\mathbf{M}'\mathbf{x}_t = \mathbf{x}_{t-1}$ 。权衡成本收益让注意力逐渐更有效体现在 \mathbf{H}'_p 与 \mathbf{H}' 的关系上, Afrouzi and Yang (2020) 编写的 DRIPs 软件包套装化求解程序而使问题得到简化。¹⁹又令

$$\begin{cases} \tilde{y}_t^e = \mathbf{H}'_y \mathbf{x}_t, \\ \tilde{z}_t^e = \mathbf{H}'_z \mathbf{x}_t. \end{cases}$$

将上述表达式代入动态理性疏忽系统的三方方程, 用待定系数法求解该系统。首先代入定价方程 $\hat{p}_{it} = \mathbb{E}_{it}^\times (\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z \tilde{z}_t^e)$, 即:

$$\begin{aligned}\mathbf{H}'_1 \tilde{\mathbf{x}}_t &= \mathbf{H}'_p \tilde{\mathbf{x}}_t + \alpha_y \mathbf{H}'_y \mathbf{x}_t + \alpha_z \mathbf{H}'_z \mathbf{x}_t, \\ \therefore \mathbf{H}'_1 \mathbf{x}_t &= \mathbf{H}'_p \mathbf{x}_t + \alpha_y \mathbf{H}'_y (\mathbf{I} - \mathbf{A}\mathbf{M}') \mathbf{x}_t + \alpha_z \mathbf{H}'_z \mathbf{x}_t,\end{aligned}$$

¹⁷https://afrouzi.com/DRIPs.jl/dev/examples/ex5_ay2020/ex5_Afrouzi_Yang_2019/

¹⁸ 参看 Hamilton (1994) p.29, 式 [2.2.8]。

¹⁹<https://afrouzi.com/DRIPs.jl/dev/>

$$\therefore \mathbf{H}_1 = \mathbf{H}_p + \alpha_y(\mathbf{I} - \mathbf{M}\mathbf{A}')\mathbf{H}_y + \alpha_z\mathbf{H}_z. \quad (3.1)$$

代入动态 IS 曲线变为 $\tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma}(\tilde{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}) + u_t$ ，即：

$$\begin{aligned} \sigma \mathbf{H}'_y \mathbf{x}_t &= \sigma \mathbf{H}'_y \mathbf{x}_{t+1} - (\mathbf{H}'_i \mathbf{x}_t - \mathbf{H}'_\pi \mathbf{x}_{t+1}) + \sigma \mathbf{H}'_u \mathbf{x}_t, \\ \therefore \mathbf{H}'_i \mathbf{x}_t &= \sigma(\mathbf{H}'_y \mathbf{M} \mathbf{x}_t - \mathbf{H}'_y \mathbf{x}_t) + \mathbf{H}'_\pi \mathbf{M} \mathbf{x}_t + \sigma \mathbf{H}'_u \mathbf{x}_t, \\ \therefore \mathbf{H}_i &= \sigma(\mathbf{M}' - \mathbf{I})\mathbf{H}_y + \mathbf{M}'\mathbf{H}_\pi + \sigma \mathbf{H}_u. \end{aligned} \quad (3.2)$$

代入利率规则 $\hat{i}_t = \rho \hat{i}_{t-1} + (1 - \rho)(\phi_\pi \hat{\pi}_t + \phi_y \tilde{y}_t) + v_t$ ，即：

$$\begin{aligned} \mathbf{H}'_i \mathbf{x}_t &= \rho \mathbf{H}'_i \mathbf{x}_{t-1} + (1 - \rho)(\phi_\pi \mathbf{H}'_\pi \mathbf{x}_t + \phi_y \mathbf{H}'_y \mathbf{x}_t) + \mathbf{H}'_v \mathbf{x}_t, \\ \therefore \mathbf{H}'_i \mathbf{x}_t - \rho \mathbf{H}'_i \mathbf{M}' \mathbf{x}_t &= (1 - \rho)\phi_\pi \mathbf{H}'_\pi \mathbf{x}_t + (1 - \rho)\phi_y \mathbf{H}'_y \mathbf{x}_t + \mathbf{H}'_v \mathbf{x}_t, \\ \therefore (\mathbf{I} - \rho \mathbf{M})\mathbf{H}_i &= (1 - \rho)\phi_\pi \mathbf{H}_\pi + (1 - \rho)\phi_y \mathbf{H}_y + \mathbf{H}_v. \end{aligned} \quad (3.3)$$

将 (3.2) 代入 (3.3)，合并整理为：

$$[\sigma(\mathbf{I} - \rho \mathbf{M})(\mathbf{M}' - \mathbf{I}) - (1 - \rho)\phi_y \mathbf{I}]\mathbf{H}_y = [(1 - \rho)\phi_\pi \mathbf{I} - (\mathbf{I} - \rho \mathbf{M})\mathbf{M}']\mathbf{H}_\pi - \sigma(\mathbf{I} - \rho \mathbf{M})\mathbf{H}_u + \mathbf{H}_v. \quad (3.4)$$

将 $\mathbf{H}_p = \mathbf{X}'\mathbf{H}_0$ (参看 Afrouzi and Yang, 2020)^{p.57} 与 $\mathbf{H}_\pi = (\mathbf{I} - \mathbf{M})(\mathbf{I} - \mathbf{M}\mathbf{A}')^{-1}\mathbf{H}_p$ 代入上式，求解 \mathbf{H}_y 后代入式 (3.1)。给定初始变量 $\mathbf{H}'_0 = \mathbf{H}'_z + \mathbf{H}'_u + \mathbf{H}'_v$ ，对式 (3.1) 迭代可求解 \mathbf{H} 。

粘性信息系统（表3右侧）中第一个方程容易推导出带有成本加成冲击项的粘性信息菲利普斯曲线 $\hat{\pi}_t = \kappa \sum_{h=0}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-1-h}(\hat{\pi}_t + \alpha_y \Delta \tilde{y}_t^e + \alpha_z \Delta z_t) + \frac{\kappa}{1 - \kappa}(\alpha_y \tilde{y}^e + \alpha_z z_t)$ 。²⁰该系统由于存在无数个滞后预期生成的状态变量，因此解法上相对复杂，但经过十数年的发展，目前求解技术已日趋成熟。²¹

下面数值模拟并比较理性疏忽（下图用上标 RI 表示）和粘性信息（用 SI 表示）的不同理论背景下发生成本加成冲击、需求冲击以及货币政策冲击后通货膨胀和产出缺口这两个宏观经济变量的动态路径。参数校准见表4。

表 4: 理性疏忽 (RI) 模型和粘性信息 (SI) 模型的参数校准

ω	β	σ	ν	γ	ϕ_π	ϕ_y	κ	ρ	ρ_z	ρ_u	ρ_v	σ_z	σ_u	σ_v
0.5	0.99	2.5	2.5	10	2	0.5/2	0.25	0.77	0.75	0.75	0.75	10	0.25	0.25

¹ ω 是信息成本参数，在表1第一行第一列信息成本参数假设为 1，现在假设 $\omega = 0.5$ ，根据 Afrouzi and Yang (2020)，该项调整为 $0.5(\log \sigma_m^2|_{t-1} - \log \sigma_m^2|_t)$ ；以下几个参数参看附录A： β 为主观贴现率， σ 为相对风险厌恶系数， ν 表示边际负效用弹性或针对劳动的相对风险规避系数， γ 是不同消费品间的替代弹性； ϕ_π 和 ϕ_y 分别是利率规则中对通货膨胀和产出缺口的政策反应参数； κ 是粘性信息理论中当期更新信息的概率，也是理性疏忽学说中的卡尔曼增益； ρ 是利率规则中利率的滞后系数； ρ_z, ρ_u, ρ_v 分别是成本加成冲击、总需求冲击、货币政策冲击的一阶滞后系数， $\sigma_z, \sigma_u, \sigma_v$ 是对应的标准差。

由附录A可知， $\alpha_y = \frac{\sigma + \nu}{1 + (1 + \frac{1}{Z})\nu}$ ， $\alpha_z = \frac{\frac{Z}{1+Z}}{1 + (1 + \frac{1}{Z})\nu}$ ， $\alpha_a = \frac{1 + \nu}{1 + (1 + \frac{1}{Z})\nu}$ ， $Z = \ln \frac{\gamma}{\gamma - 1}$ 。参考 Arslan (2008); Trabandt (2009); Afrouzi et al. (2020) 等文献，家庭主观贴现率 $\beta = 0.99$ ，令相对风险厌恶系数 $\sigma = 2.5$ ；对通货膨胀和产出缺口的政策反应系数分别是 $\phi_\pi = 2$ 和 $\phi_y = 0.25$ ；冲击的惯性系数 ρ_z, ρ_u, ρ_v 都为

²⁰ 推导过程可借鉴 Ball et al. (2005)，细节上差异的由来可参看 Paciello and Wiederholt (2014) 和 Maćkowiak and Wiederholt (2009)。

²¹ 参看 Mankiw et al. (2002); Gregory et al. (2007); Trabandt (2009); Wang et al. (2007); Arslan (2008, 2013); Menz et al. (2009); Meyer-Gohde (2010); Verona et al. (2014)。

0.75，利率规则中的惯性系数 $\rho = 0.77$ ，总需求冲击和货币政策冲击的标准差 σ_u, σ_v 皆为 0.25，为与两项冲击在相似的尺度需放大成本加成冲击，因此令 $\sigma_z = 10$ ；Mankiw and Reis (2002) 设粘性信息的概率为 0.75，此处 κ 表示拥有完美信息的比重，故为 0.25。

如图2-4所示，从单个冲击来看，在理性疏忽和粘性信息两个系统中，宏观内生变量的惯性都有较好体现。尤其，货币政策冲击之下通货膨胀的脉冲响应皆呈“驼峰状”，即宽松的货币政策实施后，高启的通货膨胀将继续升高一定时间后才回复到初始均衡状态。²²虽然两大系统的脉冲响应相似，但生成机理截然不同：以货币政策冲击为例，理性疏忽模型中，企业面临有关带有噪音的货币政冲击的信号，处理信号需要时间而导致定价继而通货膨胀的迟滞；而粘性信息模型中，迟滞是由于决策时刻未能及时获得无噪音的货币政策信号。图5显示的是三个冲击同时作用于两大系统后宏观经济变量的动态路径，从中可见货币政策冲击后通货膨胀的“驼峰”状不如单个冲击下明显，这是由于三个冲击相互影响所致。²³

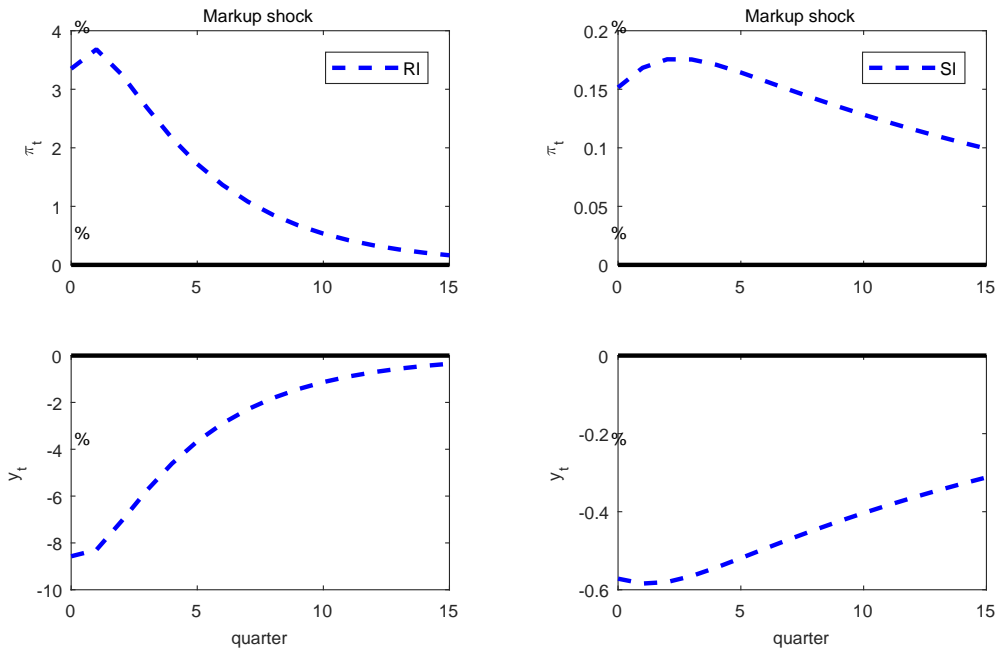


图 2: 通货膨胀（上）及产出（下）受成本加成冲击下的脉冲响应。左侧产生于理性疏忽模型（RI），右侧生成于粘性信息模型（SI），参数校准见表4。

四 最优货币政策的定量分析

至此，已初步探明粘性信息与理性疏忽的关系，即前者有条件地可成为后者的特例。区别是，前者避开信号提取，后者直面信息摩擦。如前所述，理性疏忽学说不仅在内生惯性方面拟合得更好，而且在模拟外生惯性上更符合数据特征，理论上已能确定理性疏忽更适用于货币政策分析。本节通过分别计算相对福利损失值以进一步定量比较其用于分析货币政策的优劣，这亦可提供更直接的证据。

²² 粘性价格模型被诟病的原因之一就在于未能产生“驼峰”状 (Mankiw and Reis, 2002)。

²³ 这与此处实验的单个冲击与一对冲击的不同脉冲响应是相似原理: https://afrouzi.com/DRIPs.jl/dev/examples/ex1_pricing_nofeedback/ex1_pricing_pe_nofeedback/

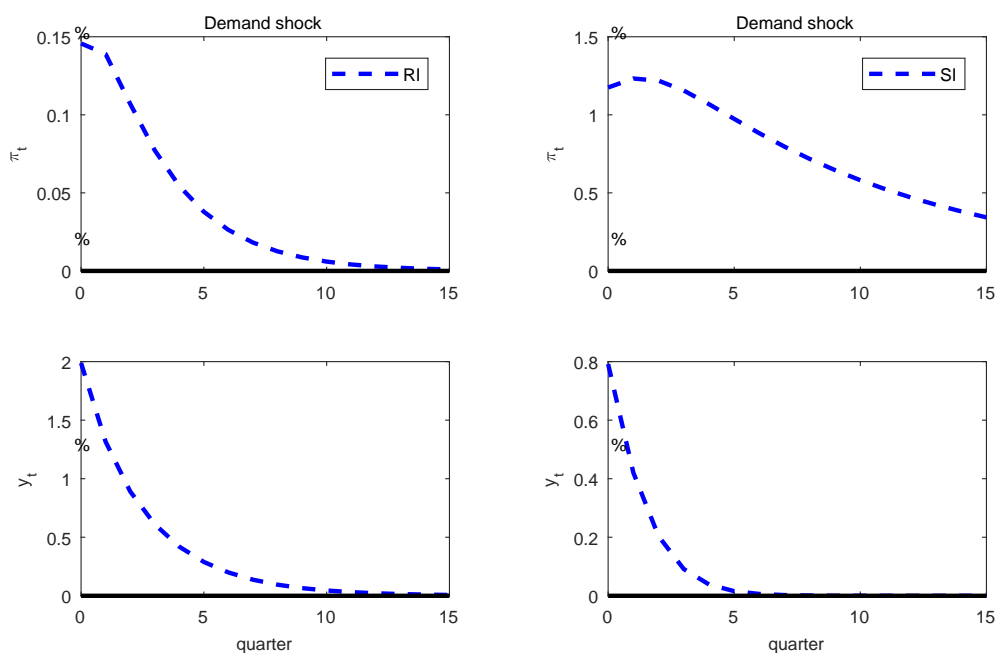


图 3: 通货膨胀（上）及产出（下）受总需求冲击下的脉冲响应。左侧产生于理性疏忽模型（RI），右侧生成于粘性信息模型（SI），参数校准见表4。

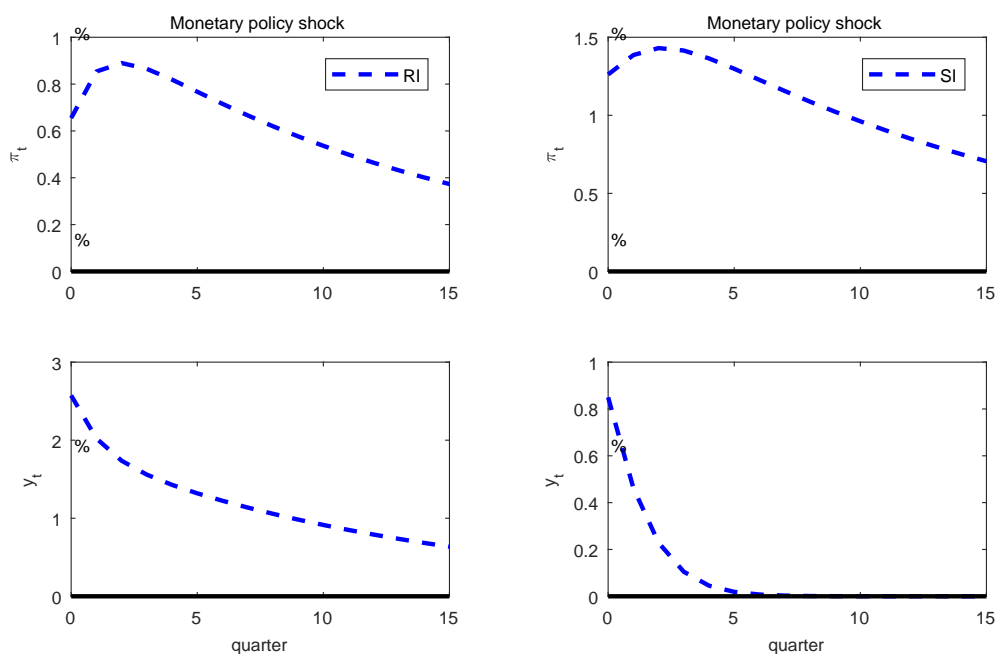


图 4: 通货膨胀（上）及产出（下）受货币政策冲击下的脉冲响应。左侧产生于理性疏忽模型（RI），右侧生成于粘性信息模型（SI），参数校准见表4。

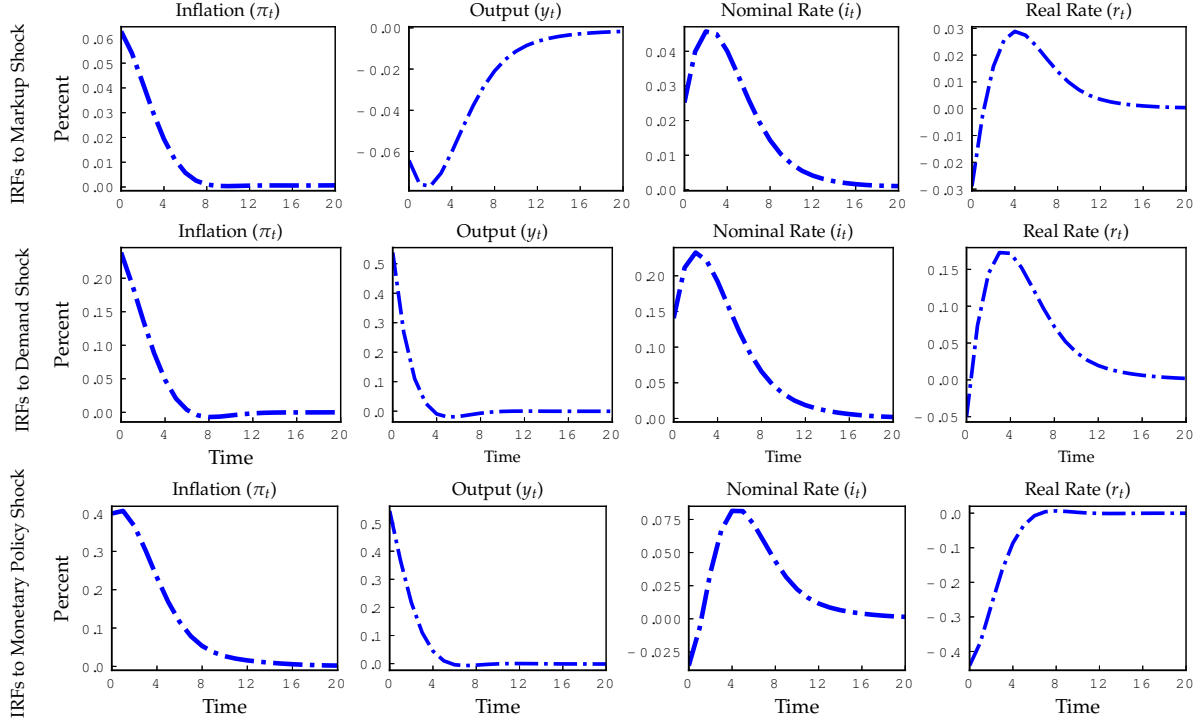


图 5: 理性疏忽模型中通货膨胀、产出、名义利率、实际利率等经济变量同时受成本加成冲击（上）、总需求冲击（中）、货币政策冲击（下）的脉冲响应。参数校准见表4。

1) 帕累托最优配置能否实现

当存在成本加成冲击时，完全信息条件下的弹性价格均衡并非有效，因此，即使央行能实现这一目标，也并非最优，最优在于能否实现有效均衡水平。通过以下分析会发现，货币政策即使不能达到有效均衡水平，至少也能达到完全信息条件下的弹性价格均衡水平。如附录A所示，产出缺口 $\tilde{y}_t = \hat{y}_t - \hat{y}_t^f$ ，其中自然产出 $\hat{y}_t^f = \frac{\alpha_a}{\alpha_y} a_t - \frac{\alpha_z}{\alpha_y} z_t$ ；Benchimol and Bounader (2019) 定义福利相关的产出缺口 $\tilde{y}_t^e = \hat{y}_t - \hat{y}_t^e$ ，其中有效产出 $\hat{y}_t^e = \frac{\alpha_a}{\alpha_y} a_t$ 。是故，产出缺口又可改写为 $\tilde{y}_t = \tilde{y}_t^e + (\hat{y}_t^e - \hat{y}_t^f)$ 。

简便起见，考察货币供给规则，分析结论适用于利率规则 (Paciello and Wiederholt, 2014)^{p.382}。回顾表1所示的局部均衡系统：

$$\begin{cases} \hat{p}_{it}^* = \hat{p}_t + \alpha_y \hat{y}_t + \alpha_z z_t - \alpha_a a_t, \\ m_t = \hat{p}_t + \hat{y}_t. \end{cases}$$

供需结合：

$$\hat{p}_{it}^* = (1 - \alpha_y) \hat{p}_t + \alpha_y \left(m_t + \frac{\alpha_z}{\alpha_y} z_t - \frac{\alpha_a}{\alpha_y} a_t \right).$$

若令货币供给规则 $m_t = \frac{\alpha_a}{\alpha_y} a_t - \frac{\alpha_z}{\alpha_y} z_t$ ，则 $\hat{p}_{it}^* = (1 - \alpha_y) \hat{p}_t$ ，进而

$$\hat{y}_t = m_t - \hat{p}_t = \frac{\alpha_a}{\alpha_y} a_t - \frac{\alpha_z}{\alpha_y} z_t - \hat{p}_t = \hat{y}_t^f - \hat{p}_t = \hat{y}_t^f.$$

显见，对于含有信息摩擦的理性疏忽或者粘性信息系统，当遭遇成本加成等冲击时，央行能实现完

全信息状态下的弹性价格均衡，虽非最优货币政策。²⁴

2) 可操作的利率规则

Paciello and Wiederholt (2014) 重点讨论了理性疏忽不同信息结构中（噪音的内、外生）基于最小化福利损失所得到的最优简单货币供给规则，本文转向研究最优简单利率规则。最优简单规则的福利损失（表5下半部分）更靠近 Ramsey 问题（表5上半部分）下的福利损失，是谓相对福利损失更小 (Huang and Liu, 2005) pp.1453-1454。下面计算并比较两大系统相对福利损失的大小，以从该视角审视其用于可操作的利率规则的优劣。简便计，暂考虑经济系统中只存在成本加成冲击。线性系统即为表3，福利损失函数的推导细节见附录C。

表 5: 最优货币政策与最优简单规则

Ramsey 问题			
理性疏忽		粘性信息	
$\min_{\tilde{y}_t^e, \hat{\pi}_t, \hat{i}_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(\tilde{y}_t^e)^2 - \frac{\tilde{y}_t^e}{\gamma} + \lambda^{\text{par}} \text{var}_i(\hat{p}_{it} - \hat{p}_t) \right],$		$\min_{\tilde{y}_t^e, \hat{\pi}_t, \hat{i}_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(\tilde{y}_t^e)^2 - \frac{\tilde{y}_t^e}{\gamma} + \lambda^{\text{par}} \text{var}_i(\hat{p}_{it} - \hat{p}_t) \right]$	
s.t. $\hat{p}_{it} = \mathbb{E}_{it}^{\times}(\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t),$		s.t. $\hat{p}_t = \kappa \sum_{h=0}^{\infty} (1-\kappa)^h \mathbb{E}_{t-h}(\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t),$	
$\tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma}(\hat{i}_t - \hat{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}).$		$\tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma}(\hat{i}_t - \hat{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}).$	
最优简单规则			
理性疏忽		粘性信息	
$\min_{\rho, \phi_{\pi}, \phi_y} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(\tilde{y}_t^e)^2 - \frac{\tilde{y}_t^e}{\gamma} + \lambda^{\text{par}} \text{var}_i(\hat{p}_{it} - \hat{p}_t) \right],$		$\min_{\rho, \phi_{\pi}, \phi_y} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(\tilde{y}_t^e)^2 - \frac{\tilde{y}_t^e}{\gamma} + \lambda^{\text{par}} \text{var}_i(\hat{p}_{it} - \hat{p}_t) \right],$	
s.t. $\hat{p}_{it} = \mathbb{E}_{it}^{\times}(\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t),$		s.t. $\hat{p}_t = \kappa \sum_{h=0}^{\infty} (1-\kappa)^h \mathbb{E}_{t-h}(\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t),$	
$\tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma}(\hat{i}_t - \hat{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}),$		$\tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma}(\hat{i}_t - \hat{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}),$	
$\hat{i}_t = \hat{i}_t^e - \rho \hat{i}_{t-1}^e + \rho \hat{i}_{t-1} + (1-\rho)(\phi_{\pi} \hat{\pi}_t + \phi_y \tilde{y}_t^e).$		$\hat{i}_t = \hat{i}_t^e - \rho \hat{i}_{t-1}^e + \rho \hat{i}_{t-1} + (1-\rho)(\phi_{\pi} \hat{\pi}_t + \phi_y \tilde{y}_t^e).$	

¹ 本表分上下两个板块，左右两列分别是理性疏忽和粘性信息背景下分析 Ramsey 问题及最优简单规则的闭合系统。Ramsey 问题的方程板块共三行，第一行是目标最小化的社会福利损失函数，第二、三行是作为约束条件的供需两端构成的完整经济系统；最优简单规则的方程板块共四行，除最后一行新增了利率规则外，其他三行不变。

由于粘性信息和理性疏忽不同的信息结构，将采用稍有不同的方法求解。对于理性疏忽系统，若噪音波动外生，则技术关键是信号提取，若噪音波动内生，则还要先求解最优注意力，Miao et al. (2019) 提出的三步法求解 Ramsey 问题更为准确；上一节介绍的 Afrouzi and Yang (2020) 的三方解法用于求解最优简单规则亦是简便。对于粘性信息系统，求解方法相对直接，Ramsey 问题而言，由目标函数和总价方程的一阶条件可确定价格与产出的最优关系，再与总价方程结合可求解目标函数中所需的变量，如有需要，根据动态 IS 曲线得到最优利率规则；最优简单规则的求解方法可结合 Dynare 实现，或者给定利率规则求解三方方程后代入目标函数，通过 Matlab 编程，确定福利损失值最小时的政策系数。²⁵

下面分别详述理性疏忽和粘性信息两大系统对于成本加成冲击的最优货币政策及最优简单规则。

(1) 理性疏忽系统的最优货币政策

根据定义， $\text{var}_i(\hat{p}_{it} - \hat{p}_t) = \int_{i=0}^{\infty} (\hat{p}_{it} - \hat{p}_t)^2 di, i \in [0, 1]$ 。要采用 Miao et al. (2019) 的三步法求解最优货币政策（Ramsey 问题），先要转换成多元线性二次型，为此先简要勾勒如何得到多元二次型的福利损失函数。²⁶

根据 Uhlig (1999) 方法，家庭部门的效用函数稍作调整为 $U(Y_t, \int_0^1 N_{it} di) = U(Y e^{\hat{y}_t}, N \int_0^1 e^{\hat{n}_{it}} di) =$

²⁴ Paciello and Wiederholt (2014) p.368 讨论的是理性疏忽系统，此处表明对于粘性信息系统也是如此。

²⁵ 作者感谢邵然慷慨分享了求解理性疏忽系统以及 Arslan 求解粘性信息系统的可供借鉴的 Matlab 代码。

²⁶ Paciello and Wiederholt (2014) 在线附录 A 及 Maćkowiak and Wiederholt (2009) 附录 A 或 Zhang (2014) 附录 B 值得借鉴。

$\tilde{u}(\hat{y}_t, \int_0^1 \hat{n}_{it} di) = \tilde{u}(\hat{y}_t, \int_0^1 (\hat{y}_{it} - a_t) di) = \tilde{u}(\hat{y}_t, \hat{y}_{it}, a_t, z_t) = \tilde{u}(\vec{y}_t, \vec{x}_t)$ 。 z_t 源自 Y_t 对 Y_{it} 的 Dixit-Stiglitz 加总。 \tilde{u} 表示二阶泰勒近似，其余变量与前文及附录的定义一致。前定状态变量 $\vec{y}_t \equiv (\hat{y}_t, \hat{y}_{it})'$ ，非前定状态变量 $\vec{x}_t \equiv (a_t, z_t)'$ 。变换后的效用函数的二次近似为：

$$\begin{aligned}\tilde{u}(\vec{y}_t, \vec{x}_t) &= \tilde{u}(0, 0) + h'_y \vec{y}_t + h'_x \vec{x}_t + \vec{y}_t' H_y \vec{y}_t + \vec{y}_t' H_{yx} \vec{x}_t + \vec{x}_t' H_x \vec{x}_t, \\ \vec{y}_t^c &\equiv \frac{\partial \tilde{u}(\vec{y}_t, \vec{x}_t)}{\partial \vec{y}_t} = 0 = h'_y + 2H_y \vec{y}_t + H_{yx} \vec{x}_t, \\ \tilde{u}(\vec{y}_t^c, \vec{x}_t) &= \tilde{u}(0, 0) + h'_y \vec{y}_t^c + h'_x \vec{x}_t + \vec{y}_t^{c'} H_y \vec{y}_t^c + \vec{y}_t^{c'} H_{yx} \vec{x}_t + \vec{x}_t' H_x \vec{x}_t,\end{aligned}$$

其中 \hat{y}_t^c 被定义为冲击后能使家庭部门效用最大化的有效产出。²⁷ $\tilde{u}(\vec{y}_t, \vec{x}_t) - \tilde{u}(\vec{y}_t^c, \vec{x}_t)$ 为冲击后的即期福利损失函数：

$$L_t = (\vec{y}_t - \vec{y}_t^c)' H_y (\vec{y}_t - \vec{y}_t^c).$$

加入控制变量 \hat{i}_t ，且不考虑技术冲击，将上述损失函数的形式改写为（视 \vec{y}_t^c 为常数纳入新的矩阵 H 中）：

$$L_t = \begin{bmatrix} z_t & \vec{y}_t & \hat{i}_t \end{bmatrix} H \begin{bmatrix} z_t \\ \vec{y}_t \\ \hat{i}_t \end{bmatrix}.$$

结合对数线性化后 \hat{y}_{it} 的需求函数及总价格水平与各商品价格水平之间的对数线性关系，并考虑成本加成冲击导致的均衡扭曲，上述多变量二次型的福利损失函数即与表5一致。如前所述，假设成本加成冲击 z_t 服从 AR(1) 过程。猜想 $\vec{y}_t = M z_t, \hat{i}_t = -F z_t$ ，Ramsey 问题转变成求解如下 Bellman 方程：

$$\begin{aligned}V(z_t) &= \min_{\hat{i}_t} L_t + \beta \mathbb{E}_t V(z_{t+1}), \\ D \begin{bmatrix} z_{t+1} \\ \mathbb{E}_t \vec{y}_{t+1} \end{bmatrix} &= A \begin{bmatrix} z_t \\ \vec{y}_t \end{bmatrix} + B \hat{i}_t + \begin{bmatrix} C \\ 0 \end{bmatrix} \epsilon_t^z.\end{aligned}$$

Miao et al. (2019) 的三步法分别是：第一步，基于完全信息求解上述 LQG 系统；第二步，利用分离原则及确定性等价原则求解外生信息结构；第三步，基于内生信息结构求解内外循环系统，求解细节不再赘述，第一、二步可另外参看Miao (2014)^{ch.9, p.297, pp.304-315}。由求解 Ramsey 问题得到的基准福利损失见表6。

最优政策规则的求解可分两步：第一步，给定政策规则，可仍借助于Afrouzi and Yang (2020) 的三方程解法得到内生变量的解；第二步，代入目标函数，选择能使福利损失最小的政策参数即为所求，同列于表6。

（2）粘性信息系统的最优货币政策

根据Ball et al. (2005) 的引理 1，可知价格偏离可进一步改写为 $\text{var}_i(\hat{p}_{it} - \hat{p}_t) = \sum_{h=1}^{\infty} \kappa_h (\hat{p}_t - \mathbb{E}_{t-h} \hat{p}_t)^2$ ，其中参数 $\kappa_h = \frac{\kappa(1-\kappa)^h}{[1-(1-\kappa)^h][1-(1-\kappa)^{h+1}]}$ 。借鉴Arslan (2013)，下面简要勾勒求解步骤。²⁸

²⁷ 常数 $\frac{1}{2}$ 已移入系数矩阵，这与Paciello and Wiederholt (2014) 在线附录 A 稍有不同。

²⁸ 成本加成冲击之下，均衡被扭曲，Arslan (2013); Ball et al. (2005) 对此稍有忽略。当然，结果将表明，稳态扭曲的出现，并不影响产出缺口和通货膨胀对冲击的响应，但会影响整个波动期间的产出缺口和通货膨胀的平均水平。当央行意欲使产出缺口为正（产出水平高于潜在产出）时，平均通货膨胀水平也将上升，从而产生通胀偏差。

构造拉格朗日函数，

$$\begin{aligned} & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(\tilde{y}_t^e)^2 - \frac{\tilde{y}_t^e}{\gamma} + \lambda^{\text{par}} \sum_{h=1}^{\infty} \kappa_h (\hat{p}_t - \mathbb{E}_{t-h} \hat{p}_t)^2 \right] \\ & + 2\mathcal{L}_t \left[\hat{p}_t - \frac{\kappa}{1-\kappa} (\alpha_y \tilde{y}_t^e + \alpha_z z_t) - \kappa \sum_{h=0}^{\infty} (1-\kappa)^h \mathbb{E}_{t-h-1} (\hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t) \right], \end{aligned}$$

注意到总价格水平的表达式已经提出了求和算子中的第一项及合并同类项，拉格朗日乘子写成 $2\mathcal{L}_t$ 为方便后续运算。

分别对 \tilde{y}_t^e 和 \hat{p}_t 求导，并联合消掉拉格朗日乘子可得：

$$\begin{aligned} \tilde{y}_t^e &= \frac{1}{2\gamma} - \lambda^{\text{par}} \alpha_y \frac{\kappa}{1-\kappa} \sum_{h=1}^{\infty} \kappa_h (\hat{p}_t - \mathbb{E}_{t-h} \hat{p}_t), \\ &= \frac{1}{2\gamma} - \Psi \Sigma_{\kappa} \hat{p}_t + \Psi \sum_{h=1}^{\infty} \kappa_h \mathbb{E}_{t-h} \hat{p}_t, \end{aligned}$$

其中， $\Psi \equiv \lambda^{\text{par}} \alpha_y \frac{\kappa}{1-\kappa}$ ， $\Sigma_{\kappa} \equiv \sum_{h=1}^{\infty} \kappa_h$ 。

用上述最优条件替换总价格水平方程中的 \tilde{y}_t^e ，可得

$$\begin{aligned} (1 - \kappa + \kappa \alpha_y \Psi \Sigma_{\kappa}) \hat{p}_t &= \frac{\kappa \alpha_y}{2\gamma} + \kappa \alpha_z z_t + \kappa \alpha_y \Psi \sum_{i=1}^{\infty} \kappa_i \mathbb{E}_{t-i} \hat{p}_t \\ &+ \kappa \sum_{h=1}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} \left[\hat{p}_t + \alpha_y \left(\frac{1}{2\gamma} - \Psi \Sigma_{\kappa} \hat{p}_t + \Psi \sum_{i=1}^{\infty} \kappa_i \mathbb{E}_{t-i} \hat{p}_t \right) + \alpha_z z_t \right]. \end{aligned}$$

稍作整理为：

$$\begin{aligned} \frac{1 - \kappa + \kappa \alpha_y \Psi \Sigma_{\kappa}}{\kappa} \hat{p}_t - \alpha_y \Psi \sum_{i=1}^{\infty} \kappa_i \mathbb{E}_{t-i} \hat{p}_t - (1 - \alpha_y \Psi \Sigma_{\kappa}) \sum_{h=1}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} \hat{p}_t - \alpha_y \Psi \sum_{h=1}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} \sum_{i=1}^{\infty} \kappa_i \mathbb{E}_{t-i} \hat{p}_t \\ = \frac{\alpha_y}{2\gamma} (1 + \Sigma_{1-\kappa}) + \alpha_z z_t + \alpha_z \sum_{h=1}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} z_t, \end{aligned}$$

其中 $\Sigma_{1-\kappa} \equiv \sum_{h=1}^{\infty} (1 - \kappa)^h$ 。

Arslan (2013) 求解的关键一步是假设冲击服从 AR(1) 过程，而非更一般地 ARMA。同样，在求解理性疏忽系统时，Paciello and Wiederholt (2014) 也只是分别假设冲击服从白噪声过程及 AR(1) 过程。Miao et al. (2019) 对此局限有所突破。简便计，姑且仍假设 z_t 服从 AR(1) 过程，可以改写成 MA(∞)，即 $z_t = \rho_z z_{t-1} + \epsilon_t^z = \sum_{j=0}^{\infty} \rho_z^j \epsilon_{t-j}^z$ ，猜想 $\hat{p}_t = \sum_{j=0}^{\infty} \theta_j^p \epsilon_{t-j}^z$ ，一起代入上式：

$$\begin{aligned} & \underbrace{\frac{1 - \kappa + \kappa \alpha_y \Psi \Sigma_{\kappa}}{\kappa}}_{a_1} \sum_{j=0}^{\infty} \theta_j^p \epsilon_{t-j}^z - \underbrace{\alpha_y \Psi}_{a_2} \sum_{i=1}^{\infty} \kappa_i \mathbb{E}_{t-i} \sum_{j=i}^{\infty} \theta_j^p \epsilon_{t-j}^z - \underbrace{(1 - \alpha_y \Psi \Sigma_{\kappa})}_{a_3} \sum_{h=1}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} \sum_{j=h}^{\infty} \theta_j^p \epsilon_{t-j}^z \\ & - \underbrace{\alpha_y \Psi}_{a_4} \sum_{h=1}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} \sum_{i=h}^{\infty} \kappa_i \mathbb{E}_{t-i} \sum_{j=i}^{\infty} \theta_j^p \epsilon_{t-j}^z, \\ & = \frac{\alpha_y}{2\gamma} (1 + \Sigma_{1-\kappa}) + \alpha_z \sum_{j=0}^{\infty} \rho_z^j \epsilon_{t-j}^z + \alpha_z \sum_{h=1}^{\infty} (1 - \kappa)^h \mathbb{E}_{t-h} \sum_{j=h}^{\infty} \rho_z^j \epsilon_{t-j}^z. \end{aligned}$$

待定系数法解得：

$$\theta_k^p = \frac{\alpha_z \rho_z^k \sum_{i=0}^k (1-\kappa)^i}{a_1 - \left[a_2 \sum_{i=1}^k \kappa_i + a_3 \sum_{i=1}^k (1-\kappa)^i + a_4 \left(\sum_{i=1}^k \kappa_i \right) \left(\sum_{i=1}^k (1-\kappa)^i \right) \right]}.$$

令 $\tilde{y}_t^e = \sum_{j=0}^{\infty} \theta_j^y \epsilon_{t-j}^z$ 代入此前推导得到的一阶条件，同理解得：

$$\theta_k^y = \Psi \left(\sum_{i=1}^k \kappa_i - \sum_{i=1}^{\infty} \kappa_i \right) \theta_k^p.$$

当 \hat{p}_t 和 \tilde{y}_t^e 被求解后，代入动态 IS 曲线，解出 \hat{i}_t ，此即为最优货币政策的显性表达式，这并非本节之重点；代入福利损失函数，即可得到福利损失值，视为基准，算出备用，见表6。

Huang and Liu (2005) 及 Galí (2015) 皆明确指出上述最优货币政策不具操作性的原因之一在于动态 IS 曲线中的自然利率不可观测，为使政策方便执行，通常施以如表5所示的简单利率规则，能使前述福利损失值最小的规则即为最优利率规则（Taylor 规则是这一类中的特例），Arslan (2013) 并没有讨论这一点，本节将基于数值解法找到此最优简单规则中的政策反应系数 ρ, ϕ_π, ϕ_y ，列于表6。²⁹

表 6: 理性疏忽（RI）和粘性信息（SI）模型中的福利损失

	最优简单规则				最优货币政策
	ρ	ϕ_π	ψ_y	$\mathcal{L}^{\text{relative}}$	$\mathcal{L}^{\text{Ramsey}}$
理性疏忽	0.45	1.84	0.16	1.1890	0.4754
粘性信息	0.23	1.71	0.37	1.8133	1.2365

¹ 本表报告了理性疏忽和粘性信息两大系统中最优货币政策（Ramsey 问题）下的福利损失及最优简单规则下的政策反应参数和相对福利损失，其中： ρ 是利率滞后系数， ϕ_π 、 ϕ_y 分别是对通货膨胀和有效产出缺口的政策反应系数； $\mathcal{L}^{\text{Ramsey}}$ 为测算的基准福利损失， $\mathcal{L}^{\text{relative}} = \frac{\mathcal{L}^{\text{osr}}}{\mathcal{L}^{\text{Ramsey}}}$ 表示相对福利损失， \mathcal{L}^{osr} 是最优简单规则下的福利损失。

如表6所示，理性疏忽系统中，基于 Ramsey 问题测算的基准福利损失值为 0.4754，最优简单规则下的相对福利损失值为 1.189，而粘性信息系统中，这两项的值分别为 1.2365 与 1.8133。相对福利损失值越靠近 1，则最优简单规则带来的福利损失越靠近 Ramsey 问题。Ramsey 问题下的最优货币政策是理论最优，由于包含不可观测的自然利率等原因，导致其在现实中具有不可操作性，自然，希望某个理论下可操作的最优简单规则产生的福利损失尽可能靠近 Ramsey 问题中的福利损失（同样重要的是，Ramsey 问题下的绝对福利损失值也是理性疏忽模型中更小）。显见，1.189 较 1.813 更靠近 1，说明理性疏忽机制可以使可操作的最优简单规则靠近更小的理论上的福利损失，这是理性疏忽学说较粘性信息更适用于货币政策分析的直接证据。

五 小结

受全球疫情及国际政治形势影响，未来几年国内外经济环境的各种不确定性显著增大。中国人民银行行长易纲 2019 年出席在日本福冈举行的二十国集团 (G20) 财长和央行行长会议时表示，“央行的宏观

²⁹ 给定 ρ, ϕ_π, ϕ_y ，此时三个约束方程三个内生变量，可直接求解后代入福利损失函数，使该目标值最小的 ρ, ϕ_π, ϕ_y 即为所求。基于 Verona and Wolters (2014)，用 Dynare 的 osr 命令亦容易测算政策反应系数及相应的福利损失值。

政策空间充足,有能力应对各种不确定性”(易纲, 2019)。中国人民银行党委书记、银保监会主席郭树清在第十二届陆家嘴论坛上的发言时也指出要通过“强化政策性金融在逆周期调节中的作用”等措施“加快恢复国民经济正常循环”(郭树清, 2020)。5G 时代已经来临,不确定性增大、逆周期形成的原因并非家庭、企业等微观经济主体不能及时获取信息,而在于信息处理能力受限。学界对信息摩擦的研究由来已久,从“信息不对称”,到“信号提取”,再到“以一定概率获得完全信息”,集大成者是“理性疏忽”,信息处理的成本收益及内生信息结构等问题在该理论中有良好定义和妥善处理,该理论用于指导制订的货币政策规则将比目前主要以粘性价格、粘性信息等为特征的新凯恩斯货币模型能更有助于国民经济实现正常循环。为支撑该论断,本文从演绎与归纳两个角度进行了论证,也相继做了定性分析与定量比较。

粘性价格和粘性信息仍是国内更为熟悉和常用的新凯恩斯理论机制,国际学术界对两者之间孰优孰劣的问题也未停止争论。粘性价格假设决策信息完美,粘性信息假设决策信息以一定概率完美。理性疏忽学说直面信息不完美,且较信号提取还能权衡信息处理的成本收益以使经济主体的注意力实现最优配置。鉴于文献中已经理清了粘性价格与理性疏忽间的关系(噪音方差与冲击方差之比等于粘性概率与非粘性概率平方之时“理性疏忽”退变为“粘性价格”),本文通过将已有的粘性信息菲利普斯曲线与演绎推导得到的理性疏忽菲利普斯曲线两相比较,进一步明确了粘性信息作为理性疏忽特例的条件:更新信息的概率等于卡尔曼增益。

通过本文前言部分的文献梳理知道,从模拟外生惯性和测算福利损的角度来看,粘性信息优于粘性价格,但在拟合内生惯性方面粘性价格模型仍优于粘性信息模型。为此,本文从拟合内生惯性、模拟外生惯性和测算福利损失入手,对粘性信息与理性疏忽做了这三方面的比较研究,结果一致显示“理性疏忽”较“粘性信息”有绝对优势,这意味着理性疏忽机制能更准确地模拟经济主体的行为特征,对中国经济波动的幅度和周期也能更好呈现,更适合用于货币政策分析,尤其在可操作层面而言,理性疏忽下的最优简单规则将会产生更小的福利损失,会更接近理想的最优货币政策。当然,这仍主要是理论层面的讨论,诚如马勇等(2017)指出的,在一些复杂条件下,特别是涉及到多种政策的协调和叠加问题时,中央银行究竟怎么操作才是合理的,“恐怕不只是一个单纯的理论问题,还需要在实践中不断探索、总结和完善。”

参 考 文 献

- 习近平, 2020. 中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议[J/OL]. 人民网. <http://jhsjk.people.cn/article/31917783>.
- 于泽, 2009. 理性非注意、粘性信息和最优货币政策[M]. 北京: 中国人民大学出版社.
- 何启志, 姚梦雨, 2017. 中国通胀预期测度及时变系数的菲利普斯曲线[J]. 管理世界(5):66-78.
- 卞志村, 胡恒强, 2016. 粘性价格、粘性信息与中国菲利普斯曲线[J]. 世界经济(4):22-43.
- 娄峰, 2016. 中国企业价格刚性研究: 基于扩展的双粘性菲利普斯曲线[J]. 中国工业经济(2):37-51.
- 张成思, 党超, 2016. 基于央行调查数据的通胀预期转化: 算法基础与理解分歧[J]. 金融评论(1):1-12.
- 彭兴韵, 2011. 粘性信息经济学——宏观经济学最新发展的一个文献综述[J]. 经济研究(12):138-151.

-
- 易纲, 2019. 中国人民银行行长易纲出席二十国集团财长和央行行长会议[J/OL]. 中国人民银行. <http://www.pbc.gov.cn/goutongjiaoliu/113456/113469/3840993/index.html>.
- 易纲, 2020. 中国人民银行行长易纲在 2020 金融街论坛上的讲话[J/OL]. 中国人民银行. <http://www.pbc.gov.cn/hanglingdao/128697/128728/128829/4113425/index.html>.
- 李拉亚, 1993. 具有理性预期的非线性系统的行为分析[J]. 经济研究(02):38-40.
- 李拉亚, 2011. 理性疏忽、粘性信息和粘性预期理论评介[J]. 经济学动态(02):119-126.
- 王军, 2009. 新凯恩斯主义粘性信息理论述评[J]. 管理世界(8):157-162.
- 王军, 2011. 克里斯托夫·西姆斯理性疏忽理论评介[J]. 经济学动态, 000(012):104-109.
- 王军, 丁玲, 2013. 理性疏忽的建模思想及其对 RBC 模型的发展[J]. 经济学动态(1):106-112.
- 肖争艳, 唐寿宁, 石冬, 2005. 中国通货膨胀预期异质性研究[J]. 金融研究(9):51-62.
- 范从来, 高洁超, 2016. 适应性学习与中国通货膨胀非均衡分析[J]. 经济研究(9):17-28.
- 董丰, 许志伟, 2020. 刚性泡沫: 基于金融风险与刚性兑付的动态一般均衡分析[J]. 经济研究(10):72-88.
- 许志伟, 樊海潮, 薛鹤翔, 2015. 公众预期、货币供给与通货膨胀动态——新凯恩斯框架下的异质性预期及其影响[J]. 经济学 (季刊), 14(4):1211-1234.
- 费兆奇, 刘康, 2020. 金融开放条件下国债市场的波动溢出和风险定价研究[J]. 经济研究(9):25-41.
- 邓燕飞, 2018. 通货膨胀动态的理论和实证研究[D]. 上海: 华东师范大学.
- 邓燕飞, 董丰, 徐迎风, 等, 2017. 价格刚性、异质性预期和通货膨胀动态[J]. 管理世界(9):17-26.
- 邓燕飞, 董丰, 张军, 2021. 垂直生产链、粘性信息与货币政策[J]. 经济学 (季刊)(已录用).
- 郭树清, 2020. 加快恢复国民经济正常循环继续深化金融体系改革开放[J/OL]. 中国人民银行. <http://www.pbc.gov.cn/hanglingdao/128697/3506068/3506301/4042017/index.html>.
- 钟春平, 田敏, 2015. 预期, 有偏性预期及其形成机制: 宏观经济学的进展与争议[J]. 经济研究(05):164-179.
- 陈彦斌, 2008. 中国新凯恩斯菲利普斯曲线研究[J]. 经济研究(12):50-64.
- 陈雨露, 2015. 重建宏观经济学的“金融支柱”[J]. 国际金融研究(6):3-11.
- 陈雨露, 马勇, 2013. 大金融论纲[M]. 北京: 中国人民大学出版社.
- 马勇, 张靖岚, 陈雨露, 2017. 金融周期与货币政策[J]. 金融研究(3):33-53.
- 齐鹰飞, 2011. 短期通货膨胀动态: 理论和中国实证[D]. 辽宁大连: 东北财经大学.
- AFROUZI H, 2020. Strategic inattention, inflation dynamics, and the non-neutrality of money[J].
- AFROUZI H, YANG C, 2018. Dynamic inattention, the phillips curve, and forward guidance[J]. IDEAS Working Paper Series from RePEc.

-
- AFROUZI H, YANG C, 2020. Dynamic rational inattention and the phillips curve[J]. Working Paper.
- ANGELETOS G M, LA'O J, 2020. Optimal monetary policy with informational frictions[J]. *Journal of Political Economy*, 128(3):1027-1064.
- ARSLAN M M, 2008. Dynamics of sticky information and sticky price models in a new keynesian dsge framework[J]. *Economic Modelling*, 25:1276-1294.
- ARSLAN M M, 2010. Relative importance of sticky prices and sticky information in price setting[J]. *Economic Modelling*, 27(5):1124-1135.
- ARSLAN M M, 2013. Optimal monetary policy with the sticky information model of price adjustment: Inflation or price-level targeting?[J]. *Bulletin of Economic Research*, 65(s1):s106-s129.
- BALL L, MANKIW N G, REIS R, 2005. Monetary policy for inattentive economies[J]. *Journal of Monetary Economics*, 52(4):703-725.
- BENCHIMOL J, BOUNADER L, 2019. Optimal monetary policy under bounded rationality[J]. *International Monetary Fund*, WP19(166).
- CALVO G A, 1983. Staggered prices in a utility-maximizing framework[J]. *Journal of monetary Economics*, 12(3):383-398.
- CARRILLO J A, 2012. How well does sticky information explain the dynamics of inflation, output, and real wages[J]. *Journal of Economic Dynamics & Control*, 36:830-850.
- CHANG C, CHEN K, WAGGONER D F, et al., 2015. Trends and cycles in china's macroeconomy[J]. *Social Science Electronic Publishing*, 30(1):1-84.
- CHRISTIANO L J, EICHENBAUM M, EVANS C L, 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy[J]. *Journal of Political Economy*, 113(1):1-45.
- CHRISTIANO L J, TRABANDT M, VALENTIN K, 2010. Dsge models for monetary policy analysis [J]. *Handbook of Monetary Economics*, 3(16074):285-367.
- CLARIDA R, GALI J, GERTLER M, 1999. The science of monetary policy: A new keynesian perspective[J]. *Journal of Economic Literature*, 37:1661-1707.
- COIBION O, 2010. Testing the sticky information phillips curve[J]. *Review of Economics & Statistics*, 92(1):87-101.
- COIBION O, GORODNICHENKO Y, 2012. What can survey forecasts tell us about informational rigidities?[J]. *Journal of Political Economy*, 120(1):116-159.
- COIBION O, GORODNICHENKO Y, 2015. Information rigidity and the expectations formation process: A simple framework and new facts[J]. *American Economic Review*, 105(8):2644-2678.
- COIBION O, GORODNICHENKO Y, KAMDAR R, 2018a. The formation of expectations, inflation, and the phillips curve†[J]. *Journal of Economic Literature*, 56(4):1447-1491.

-
- COIBION O, GORODNICHENKO Y, KUMAR S, 2018b. How do firms form their expectations? new survey evidence[J]. *American Economic Review*, 108(9):2671-2713.
- COIBION O, GORODNICHENKO Y, ROPELE T, 2020. Inflation expectations and firm decisions: New causal evidence[J]. *The Quarterly Journal of Economics*, 135(1):165-219.
- DONG F, WEN Y, 2019. Long and plosser meet bewley and lucas[J]. *Journal of Monetary Economics*, 102:70-92.
- DONG F, MIAO J, WANG P, 2020. Asset bubbles and monetary policy[J]. *Review of Economic Dynamics*, 37:s68-s98.
- DUPOR B, HAN J, TSAI Y C, 2009. What do technology shocks tell us about the new keynesian paradigm[J]. *Journal of Monetary Economics*, 56(4):560-569.
- DUPOR B, KITAMURA T, TSURUGA T, 2010. Integrating sticky prices and sticky information[J]. *Review of Economics & Statistics*, 92(3):657-669.
- EGGERTSSON G B, GARGA V, 2019. Sticky prices versus sticky information: Does it matter for policy paradoxes?[J]. *Review of Economic Dynamics*, 31:363-392.
- FRIEDMAN M M, 1968. The role of monetary policy[J]. *American Economic Review*, 58(1).
- FRIEDMAN M, SCHWARTZ A J, 1963. A monetary history of the united states, 1867-1960[J]. *NBER Books*, 70(1):512-523.
- FUHRER J, 2012. The role of expectations in inflation dynamics[J]. *International Journal of Central Banking*, 8(3):137-165.
- GALI J, GERTLER M, 1999. Inflation dynamics: A structural econometric investigation[J]. *Journal of Monetary Economics*, 44:195-222.
- GALÍ J, 2015. Monetary policy, inflation, and the business cycle: An introduction to the new keynesian framework and its applications[M]. 2nd ed. Princeton, New Jersey: Princeton University Press,.
- GORDON R, 1997. The time-varying nairu and its implications for monetary policy[J]. *Journal of Economic Perspectives*, 11:11-32.
- GREGORY M N, RICARDO R, 2007. Sticky information in general equilibrium[J]. *Journal of the European Economic Association*, 5(2-3):603-613.
- HAMILTON J D, 1994. Time series analysis[M]. 1st ed. Princeton, New Jersey: Princeton University Press,.
- HÉBERT B, WOODFORD M, et al., 2018. Rational inattention in continuous time[J]. Unpublished manuscript, September.
- HUANG K X D, LIU Z, 2005. Inflation targeting: What inflation rate to target?[J]. *Journal of Monetary Economics*, 52(8):1435-1462.

-
- KEEN B D, 2007. Sticky price and sticky information price-setting models: What is the difference?[J]. *Economic Inquiry*, 45(4):770-786.
- KILEY M T, 2007. A quantitative comparison of sticky-price and sticky-information models of price setting[J]. *Journal of Money Credit & Banking*, 39(S1):101-125.
- KITAMURA T, 2008. Optimal monetary policy under sticky prices and sticky information[J]. The Ohio State University and The Bank of Japan.
- KLENOW P J, WILLIS J L, 2007. Sticky information and sticky prices[J]. *Journal of Monetary Economics*, 54:79-99.
- KNOTEK II E S, 2010. A tale of two rigidities: Sticky prices in a sticky-information environment[J]. *Journal of Money, Credit and Banking*, 42(8):1543-1564.
- KORENOK O, 2008. Empirical comparison of sticky price and sticky information models[J]. *Journal of Macroeconomics*, 30(3):906-927.
- LUCAS R E J, 1972. Expectations and the neutrality of money[J]. *Journal of Economic Theory*, 4(2): 103-124.
- LUCAS R E J, 1973. Some international evidence on output-inflation tradeoffs[J]. *American Economic Review*, 63(3):326-334.
- LUO Y, 2008. Consumption dynamics under information processing constraints[J]. *Review of Economic dynamics*, 11(2):366-385.
- LUO Y, YOUNG E R, 2013. Signal extraction and rational inattention[J]. *Economic Inquiry*, 52(2): 811-829.
- LUO Y, NIE J, YOUNG E R, 2012. Robustness, information-processing constraints, and the current account in small open economies[J]. *Journal of International Economics*, 88(1):104-120.
- LUO Y, NIE J, WANG G, et al., 2017. Rational inattention and the dynamics of consumption and wealth in general equilibrium[J]. *Journal of Economic Theory*, 172:55-87.
- MAĆKOWIAK B, WIEDERHOLT M, 2009. Optimal sticky prices under rational inattention[J]. *American Economic Review*, 99(3):769-803.
- MAĆKOWIAK B, WIEDERHOLT M, 2015. Business cycle dynamics under rational inattention[J]. *Review of Economic Studies*, 82:1502-1532.
- MAĆKOWIAK B, MATĚJKA F, WIEDERHOLT M, 2018a. Dynamic rational inattention: Analytical results[J]. *Journal of Economic Theory*, 176:650-692.
- MAĆKOWIAK B, MATĚJKA F, WIEDERHOLT M, 2018b. Survey: Rational inattention, a disciplined behavioral model[J]. *CEPR Discussion Papers*, 13243.
- MAĆKOWIAK B A, MATEJKA F, WIEDERHOLT M, 2016. The rational inattention filter[J]. *CEPR Discussion Papers*.

-
- MANKIW N G, REIS R, 2002. Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve[J]. *Quarterly Journal of Economics*, 117(4):1295-1328.
- MATEJKA F, 2016. Rationally inattentive seller: Sales and discrete pricing[J]. *The Review of Economic Studies*, 83(3):1125-1155.
- MAVROEIDIS S, PLAGBORG-MØLLER M, STOCK J H, 2014. Empirical evidence on inflation expectations in the new keynesian phillips curve[J]. *Journal of Economic Literature*, 52(1):124-188.
- MELOSI L, 2014. Estimating models with dispersed information[J]. *American Economic Journal: Macroeconomics*, 6(1):1-31.
- MENZ J O, VOGEL L, 2009. A detailed derivation of the sticky price and sticky information new keynesian dsge model[R]. [S.l.]: DEP (Socioeconomics) Discussion Papers, Macroeconomics and Finance Series.
- MEYER-GOHDE A, 2010. Linear rational-expectations models with lagged expectations: A synthetic method[J]. *Journal of Economic Dynamics & Control*, 34(5):984-1002.
- MIAO J, 2014. *Economic dynamics in discrete time*[M]. Cambridge, MA: the MIT Press.
- MIAO J, 2019. Multivariate lqg control under rational inattention in continuous time[R]. [S.l.]: Boston University-Department of Economics.
- MIAO J, WU J, YOUNG E, 2019. Multivariate rational inattention[R]. [S.l.]: Boston University-Department of Economics.
- MONDRIA J, WU T, 2010. The puzzling evolution of the home bias, information processing and financial openness[J]. *Journal of Economic Dynamics and Control*, 34(5):875-896.
- NUNES R, 2010. Inflation dynamics: The role of expectations[J]. *Journal of Money Credit & Banking*, 42(6):1161-1172.
- PACIELLO L, WIEDERHOLT M, 2014. Exogenous information, endogenous information, and optimal monetary policy[J]. *Review of Economic Studies*, 81(1):356-388.
- PACIELLO L, 2012. Monetary policy and price responsiveness to aggregate shocks under rational inattention[J]. *Journal of Money, Credit and Banking*, 44(7):1375-1399.
- PASTEN E, SCHOENLE R, 2016. Rational inattention, multi-product firms and the neutrality of money[J]. *Journal of Monetary Economics*, 80:1-16.
- PENG L, XIONG W, 2006. Investor attention, overconfidence and category learning[J]. *Journal of Financial Economics*, 80(3):563-602.
- ROTEMBERG J, 1982. Monopolistic price adjustment and aggregate output[J]. *Review of Economic Studies*, 44:517-531.
- ROTEMBERG J, WOODFORD M, 1997. An optimization-based econometric framework for the evaluation of monetary policy[J]. *NBER Macroeconomics Annual*:297-346.

-
- SIMS C A, 1998. Stickiness[C]//In Conference Series on Public Policy. USA: Carnegie-Rochester: 11/21-22/97.
- SIMS C A, 2003. Implications of rational inattention[J]. Journal of Monetary Economics, 50.
- SIMS C A, 2010. Rational inattention and monetary economics[J]. Handbook of Monetary Economics, 3.
- STAIGER D, STOCK J, WATSON M, 1997. The nairu, unemployment, and monetary policy[J]. Journal of Economic Perspectives, 11:33-51.
- TAYLOR J B, 1979. Staggered wage setting in a macro model[J]. American Economic Review, 69(2): 108-113.
- TAYLOR J B, 1980. Aggregate dynamics and staggered contracts[J]. Journal of Political Economy, 88: 1-22.
- TRABANDT M, 2009. Sticky information vs. sticky prices: A horse race in a dsge framework[J]. Working Paper Series(209).
- TUTINO A, 2013. Rationally inattentive consumption choices[J]. Review of Economic Dynamics, 16 (3):421-439.
- UHLIG H, 1999. A toolkit for analyzing nonlinear dynamic stochastic models easily[C]//MARIMON R, SCOTT A. Computational Methods for the Study of Dynamic Economies. USA: Oxford University Press: 30-62.
- VERONA F, WOLTERS M H, 2014. Sticky information models in dynare[J]. Computational Economics, 43(3):357-370.
- WANG P, WEN Y, 2007. Inflation dynamics: A cross-country investigation[J]. Journal of Monetary Economics, 54(7):2004-2031.
- WIEDERHOLT M, 2010. Rational inattention[J]. The New Palgrave Dictionary of Economics (Online Edition ed.).
- WOODFORD M, 2001. The taylor rule and optimal monetary policy[J]. American Economic Review, 91(2):232-237.
- WOODFORD M, 2002. Imperfect common knowledge and the effects of monetary policy[C]// AGHION P, FRYDMAN R, STIGLITZ J, et al. Knowledge, Information and Expectations in Modern Macroeconomics. Princeton: Princeton University Press: 195-214.
- YANG C, 2019. Rational inattention, menu costs, and multi-product firms: Micro evidence and aggregate implications[J].
- ZHANG F, 2014. Monetary policy for rationally inattentive economies with staggered price setting[J]. Journal of Economic Dynamics & Control, 38(1):184-208.
- ZORN P, 2016. Investment under rational inattention: Evidence from us sectoral data[J].

技术附录

附录 A DSGE 的基本构造

假设家庭、企业、政府部门皆有完美信息，产品市场存在垄断竞争且替代弹性时变，对数线性化后的合意定价通过以下推导步骤可得。

(1) 家庭部门

$$\begin{aligned} \max_{C_t, N_t, B_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(C_t, N_t), \\ U(C_t, N_t) = \frac{C_t^{1-\sigma} - 1}{1-\sigma} - \frac{N_t^{1+\nu}}{1+\nu}, \\ \int_0^1 P_{it} C_{it} di + Q_t B_t \leq \int_0^1 W_t N_{it} di + B_{t-1} + D_t - T_t, \\ C_t = \left(\int_0^1 C_{it}^{\frac{\gamma_t-1}{\gamma_t}} di \right)^{\frac{\gamma_t}{\gamma_t-1}}. \end{aligned}$$

其中 C_t 是第 t 期的家庭消费， N_t 是家庭提供的劳动力总和， $U(\cdot)$ 是即期效用函数， β 是家庭主观贴现率， \mathbb{E}_0 是立于初期（第 0 期）对未来的预期算子， $t = 0 \rightarrow \infty$ 表示终身， σ 和 ν 分别是消费和劳动的相对风险厌恶系数， B_t 和 Q_t 分别是债券及购买价格，劳动力市场完全竞争工资尽皆相同为 W_t ， D_t 和 T_t 分别是企业给家庭的股权分红及国家对家庭的一次性赋税， γ_t 是不同消费品之间的时变替代弹性。第一个方程是目标函数，第二个是即期效用函数的具体形式，第三个是家庭预算约束，第四个是定义的一篮子消费品的构成形式。

家庭的决策分两步，首先是给定一篮子商品支出最小化，再是预算约束下效用最大化。得到后续推导所要用的方程分别为：

$$\begin{aligned} C_{it} &= \left(\frac{P_{it}}{P_t} \right)^{-\gamma_t} C_t, \\ P_t &= \left(\int_0^1 P_{it}^{1-\gamma_t} di \right)^{\frac{1}{1-\gamma_t}}, \\ \frac{W_t}{P_t} &= C_t^\sigma N_t^\nu. \end{aligned}$$

上述第一个方程是对消费品 i 的需求函数，第二个方程是总价格指数的构成形式，第三个方程是劳动力供给方程。

(2) 企业部门

$$\begin{aligned} \max_{P_{it}} \quad & \Pi_{it} \equiv (1 - \tau_p) \frac{P_{it}}{P_t} Y_{it} - \frac{W_t}{P_t} N_{it} \\ s.t. \quad & Y_{it} = A_t N_{it} \\ & Y_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\gamma_t} Y_t. \end{aligned}$$

上述为企业利润最大化的目标函数及供需约束条件， Y_{it} 和 Y_t 分别是第 t 期企业 i 的产出及部门总体产出， N_{it} 是第 t 期企业 i 使用的劳动力， τ_p 表示定价税。将约束条件代入目标函数，将有约束的最优化问题转化为无约束的最优问题：

$$\begin{aligned}\Pi_{it} &\equiv (1 - \tau_p) \frac{P_{it}}{P_t} \left(\frac{P_{it}}{P_t} \right)^{-\gamma_t} Y_t - \frac{W_t}{P_t} \left(\frac{P_{it}}{P_t} \right)^{-\gamma_t} \frac{Y_t}{A_t}, \\ &= (1 - \tau_p) Y_t \left(\frac{P_{it}}{P_t} \right)^{1-\gamma_t} - \frac{W_t}{P_t} \frac{Y_t}{A_t} \left(\frac{P_{it}}{P_t} \right)^{-\gamma_t}.\end{aligned}$$

一阶条件为：

$$\begin{aligned}(1 - \tau_p)(1 - \gamma_t) \left(\frac{P_{it}^*}{P_t} \right)^{-\gamma_t} \frac{Y_t}{P_t} &= (-\gamma_t) \frac{W_t}{P_t} \frac{1}{A_t} \frac{Y_t}{P_t} \left(\frac{P_{it}^*}{P_t} \right)^{-\gamma_t-1}, \\ \Rightarrow (1 - \tau_p)(1 - \gamma_t) &= (-\gamma_t) \frac{W_t}{P_t} \frac{1}{A_t} \left(\frac{P_{it}^*}{P_t} \right)^{-1}, \\ \Rightarrow \frac{P_{it}^*}{P_t} &= \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{W_t/P_t}{A_t} = \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} MC_t^r.\end{aligned}$$

其中 MC_t^r 表示实际边际成本，即将出现的 MRS 表示边际替代率。根据家庭部门的最优条件及均衡条件可知 $\frac{W_t}{P_t} = MRS_t = Y_t^\sigma N_t^\nu$ ，代入上式：

$$\begin{aligned}\frac{P_{it}^*}{P_t} &= \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^\sigma N_t^\nu}{A_t}, \\ &= \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^\sigma}{A_t} \left(\int_0^1 N_{it} di \right)^\nu, \\ &= \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^\sigma}{A_t} \left[\int_0^1 \left(\frac{Y_{it}}{A_t} \right) di \right]^\nu, \\ &= \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^\sigma}{A_t^{1+\nu}} \left(\int_0^1 Y_{it} di \right)^\nu, \\ &= \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^\sigma}{A_t^{1+\nu}} \left[\int_0^1 \left(\frac{P_{it}^*}{P_t} \right)^{-\gamma_t} Y_{it} di \right]^\nu, \\ &= \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^{\sigma+\nu}}{A_t^{1+\nu}} \left[\int_0^1 \left(\frac{P_{it}^*}{P_t} \right)^{-\gamma_t} di \right]^\nu, \\ &\stackrel{?}{=} \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^{\sigma+\nu}}{A_t^{1+\nu}} \left(\frac{P_{it}^*}{P_t} \right)^{-\gamma_t \nu}. \\ \because P_{it}^* &= P_t^* \quad \therefore \frac{P_{it}^*}{P_t} = \frac{P_t^*}{P_t} = \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{Y_t^{\sigma+\nu}}{A_t^{1+\nu}} \left(\frac{P_t^*}{P_t} \right)^{-\gamma_t \nu}, \\ \because P_t^* &= P_t \quad \therefore 1 = \frac{\gamma_t}{\gamma_t - 1} \frac{1}{1 - \tau_p} \frac{\left(Y_t^f \right)^{\sigma+\nu}}{A_t^{1+\nu}}.\end{aligned}$$

对数线性化， $x_t = \ln \frac{X_t}{X}$ ， $X_t = X e^{x_t}$ ，

$$\frac{\gamma_t}{\gamma_t - 1} \equiv 1 + Z_t, \quad \ln \frac{\gamma_t}{\gamma_t - 1} = \ln(1 + Z_t) \approx Z_t, \quad z_t = \ln Z_t - \ln Z; \quad \gamma_t = 1 + \frac{1}{Z_t}.$$

$$\begin{aligned}
 & \left(\frac{P_t^*}{P_t} \right)^{1+\gamma_t\nu} = \frac{\gamma_t}{\gamma_t-1} \frac{1}{1-\tau_p} \frac{Y_t^{\sigma+\nu}}{A_t^{1+\nu}}, \\
 \Rightarrow & \left(\frac{P_t^*}{P_t} \right)^{1+(1+\frac{1}{Z_t})\nu} = (1+Z_t) \frac{1}{1-\tau_p} \frac{Y_t^{\sigma+\nu}}{A_t^{1+\nu}}, \\
 \Rightarrow & \left[1 + \left(1 + \frac{1}{Z_t} \right) \nu \right] (\ln P_t^* - \ln P_t) = \ln(1+Z_t) + \ln \frac{1}{1-\tau_p} + (\sigma+\nu) \ln Y_t - (1+\nu) \ln A_t, \\
 \Rightarrow & \left[1 + \left(1 + \frac{1}{Z_t} \right) \nu \right] (\ln P_t^* - \ln P_t) = \frac{Z}{1+Z} \ln Z + \ln \frac{1}{1-\tau_p} + (\sigma+\nu) \ln Y_t - (1+\nu) \ln A_t, \\
 \Rightarrow & \left[1 + \left(1 + \frac{1}{Z} \right) \nu \right] (\ln P^* - \ln P) = \frac{Z}{1+Z} \ln Z + \ln \frac{1}{1-\tau_p} + (\sigma+\nu) \ln Y - (1+\nu) \ln A, \\
 \Rightarrow & \hat{p}_t^* - \hat{p}_t \approx \frac{\sigma+\nu}{1+(1+\frac{1}{Z})\nu} \hat{y}_t + \frac{\frac{Z}{1+Z}}{1+(1+\frac{1}{Z})\nu} z_t - \frac{1+\nu}{1+(1+\frac{1}{Z})\nu} a_t, \\
 \Rightarrow & \hat{p}_t^* = \hat{p}_t + \alpha_y \hat{y}_t + \alpha_z z_t - \alpha_a a_t,
 \end{aligned}$$

其中, $\alpha_y = \frac{\sigma+\nu}{1+(1+\frac{1}{Z})\nu}$, $\alpha_z = \frac{\frac{Z}{1+Z}}{1+(1+\frac{1}{Z})\nu}$, $\alpha_a = \frac{1+\nu}{1+(1+\frac{1}{Z})\nu}$.

$$\begin{aligned}
 & \hat{p}_t^* = \hat{p}_t + \alpha_y \hat{y}_t + \alpha_z z_t - \alpha_a a_t, \\
 \Rightarrow & \hat{y}_t = \frac{1}{\alpha_y} (\hat{p}_t^* - \hat{p}_t) + \frac{\alpha_a}{\alpha_y} a_t - \frac{\alpha_z}{\alpha_y} z_t, \\
 \Rightarrow & \hat{y}_t^f = \frac{\alpha_a}{\alpha_y} a_t - \frac{\alpha_z}{\alpha_y} z_t, \quad \hat{y}_t^e = \frac{\alpha_a}{\alpha_y} a_t. \\
 & \hat{y}_t - \hat{y}_t^f = \frac{1}{\alpha_y} (\hat{p}_t^* - \hat{p}_t), \\
 \Rightarrow & \hat{p}_t^* = \hat{p}_t + \alpha_y (\hat{y}_t - \hat{y}_t^f), \\
 \Rightarrow & \hat{p}_t^* = \hat{p}_t + \alpha_y \tilde{y}_t; \tag{A.1} \\
 & \hat{y}_t - \hat{y}_t^e = \frac{1}{\alpha_y} (\hat{p}_t^* - \hat{p}_t) - \frac{\alpha_z}{\alpha_y} z_t, \\
 \Rightarrow & \hat{p}_t^* = \hat{p}_t + \alpha_y (\hat{y}_t - \hat{y}_t^e) + \alpha_z z_t, \\
 \Rightarrow & \hat{p}_t^* = \hat{p}_t + \alpha_y \tilde{y}_t^e + \alpha_z z_t. \tag{A.2}
 \end{aligned}$$

其中 $\tilde{y}_t \equiv \hat{y}_t - \hat{y}_t^f$, $\tilde{y}_t^e \equiv \hat{y}_t - \hat{y}_t^e$ 。当不存在成本加成冲击时, $\hat{y}_t^f = \hat{y}_t^e$, 也即 $\tilde{y}_t = \tilde{y}_t^e$ 。

需求侧的动态 IS 曲线作如下简单变换:

$$\begin{aligned}
 & \hat{y}_t = \mathbb{E}_t \hat{y}_{t+1} - \frac{1}{\sigma} (\hat{i}_t - \mathbb{E}_t \hat{\pi}_{t+1}), \\
 \Rightarrow & \hat{y}_t^e = \mathbb{E}_t \hat{y}_{t+1}^e - \frac{1}{\sigma} (\hat{i}_t^e - 0), \\
 \Rightarrow & \hat{y}_t - \hat{y}_t^e = \mathbb{E}_t (\hat{y}_{t+1} - \hat{y}_{t+1}^e) - \frac{1}{\sigma} (\hat{i}_t - \hat{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}), \\
 \Rightarrow & \tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma} (\tilde{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}), \\
 \Rightarrow & \tilde{y}_t^e = \mathbb{E}_t \tilde{y}_{t+1}^e - \frac{1}{\sigma} (\tilde{i}_t^e - \mathbb{E}_t \hat{\pi}_{t+1}) + u_t,
 \end{aligned}$$

其中 $\tilde{i}_t^e \equiv \hat{i}_t - \hat{i}_t^e$; 而 $\hat{i}_t^e = \sigma \mathbb{E}_t \Delta \hat{y}_{t+1}^e = \frac{\sigma \alpha_a}{\alpha_y} \mathbb{E}_t \Delta a_{t+1}$; ³⁰ u_t 可理解为由偏好变动等导致的需求冲击, 此处省略了这一常见的推导过程。

附录 B 系数矩阵中元素的确立

$$\begin{aligned}
 & \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 \\ G & H \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ G & H \end{bmatrix}' - \begin{bmatrix} 1 & 0 \\ G & H \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}' \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sigma_\xi^2 \right)^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}' \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ G & H \end{bmatrix}' + \begin{bmatrix} 1 \\ B \end{bmatrix} \begin{bmatrix} 1 \\ B \end{bmatrix}' \sigma_\epsilon^2, \\
 &= \begin{bmatrix} 1 & 0 \\ G & H \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & G \\ 0 & H \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ G & H \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sigma_\xi^2 \right)^{-1} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & G \\ 0 & H \end{bmatrix} + \begin{bmatrix} 1 \\ B \end{bmatrix} \begin{bmatrix} 1 & B \end{bmatrix} \sigma_\epsilon^2, \\
 &= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ G\Sigma_{11} + H\Sigma_{21} & G\Sigma_{12} + H\Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & G \\ 0 & H \end{bmatrix} - \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ G\Sigma_{11} + H\Sigma_{21} & G\Sigma_{12} + H\Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left(\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sigma_\xi^2 \right)^{-1} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & G \\ 0 & H \end{bmatrix} + \begin{bmatrix} 1 & B \\ B & B^2 \end{bmatrix} \sigma_\epsilon^2, \\
 &= \begin{bmatrix} \Sigma_{11} & \Sigma_{11}G + \Sigma_{12}H \\ G\Sigma_{11} + H\Sigma_{21} & G^2\Sigma_{11} + GH\Sigma_{21} + GH\Sigma_{12} + H^2\Sigma_{22} \end{bmatrix} - \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ G\Sigma_{11} + H\Sigma_{21} & G\Sigma_{12} + H\Sigma_{22} \end{bmatrix} (\Sigma_{11} + \sigma_\xi^2)^{-1} \begin{bmatrix} \Sigma_{11} & \Sigma_{11}G + \Sigma_{12}H \end{bmatrix} + \begin{bmatrix} \sigma_\epsilon^2 & B\sigma_\epsilon^2 \\ B\sigma_\epsilon^2 & B^2\sigma_\epsilon^2 \end{bmatrix}, \\
 &= \begin{bmatrix} \Sigma_{11} & G\Sigma_{11} + H\Sigma_{12} \\ G\Sigma_{11} + H\Sigma_{21} & G^2\Sigma_{11} + GH\Sigma_{21} + GH\Sigma_{12} + H^2\Sigma_{22} \end{bmatrix} - \begin{bmatrix} \frac{\Sigma_{11}^2}{\Sigma_{11} + \sigma_\xi^2} & \frac{G\Sigma_{11}^2 + H\Sigma_{11}\Sigma_{12}}{\Sigma_{11} + \sigma_\xi^2} \\ \frac{G\Sigma_{11}^2 + H\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2} & \frac{(G\Sigma_{11} + H\Sigma_{21})(G\Sigma_{11} + H\Sigma_{12})}{\Sigma_{11} + \sigma_\xi^2} \end{bmatrix} + \begin{bmatrix} \sigma_\epsilon^2 & B\sigma_\epsilon^2 \\ B\sigma_\epsilon^2 & B^2\sigma_\epsilon^2 \end{bmatrix}, \\
 &= \begin{bmatrix} \Sigma_{11} - \frac{\Sigma_{11}^2}{\Sigma_{11} + \sigma_\xi^2} + \sigma_\epsilon^2 & G\Sigma_{11} + H\Sigma_{12} - \frac{G\Sigma_{11}^2 + H\Sigma_{11}\Sigma_{12}}{\Sigma_{11} + \sigma_\xi^2} + B\sigma_\epsilon^2 \\ G\Sigma_{11} + H\Sigma_{21} - \frac{G\Sigma_{11}^2 + H\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2} + B\sigma_\epsilon^2 & G^2\Sigma_{11} + GH\Sigma_{21} + GH\Sigma_{12} + H^2\Sigma_{22} - \frac{(G\Sigma_{11} + H\Sigma_{21})(G\Sigma_{11} + H\Sigma_{12})}{\Sigma_{11} + \sigma_\xi^2} + B^2\sigma_\epsilon^2 \end{bmatrix},
 \end{aligned}$$

将 $G = \kappa$, $H = 1 - \kappa$, $B = \kappa$ 代入后有:

$$\begin{aligned}
 \Sigma_{11} &= \Sigma_{11} - \frac{\Sigma_{11}^2}{\Sigma_{11} + \sigma_\xi^2} + \sigma_\epsilon^2; \\
 \Sigma_{21} &= G\Sigma_{11} + H\Sigma_{21} - \frac{G\Sigma_{11}^2 + H\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2} + B\sigma_\epsilon^2, \\
 &= \kappa\Sigma_{11} + (1 - \kappa)\Sigma_{21} - \frac{\kappa\Sigma_{11}^2 + (1 - \kappa)\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2} + \kappa\sigma_\epsilon^2, \\
 &= \kappa \left(\Sigma_{11} - \frac{\Sigma_{11}^2}{\Sigma_{11} + \sigma_\xi^2} \right) + (1 - \kappa)\Sigma_{21} - \frac{(1 - \kappa)\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2} + \kappa\sigma_\epsilon^2, \\
 &= \kappa \left(\Sigma_{11} - \frac{\Sigma_{11}^2}{\Sigma_{11} + \sigma_\xi^2} + \sigma_\epsilon^2 \right) + (1 - \kappa)\Sigma_{21} - \frac{(1 - \kappa)\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2}, \\
 &= \kappa\Sigma_{11} + (1 - \kappa)\Sigma_{21} - \frac{(1 - \kappa)\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2}.
 \end{aligned}$$

由于 $\Sigma_{11}, \Sigma_{21} > 0$, Σ_{11} 易解得为 $\frac{\sigma_\epsilon^2 + \sqrt{\sigma_\epsilon^4 + 4\sigma_\xi^2\sigma_\epsilon^2}}{2}$, 或变换为 $\frac{\sigma_\epsilon^2 \left(1 + \sqrt{1 + 4\frac{\sigma_\xi^2}{\sigma_\epsilon^2}} \right)}{2}$ 。

$$\begin{aligned}
 \Sigma_{21} &= \kappa\Sigma_{11} + (1 - \kappa)\Sigma_{21} - \frac{(1 - \kappa)\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2}, \\
 \Rightarrow \kappa\Sigma_{21} &= \kappa\Sigma_{11} - \frac{(1 - \kappa)\Sigma_{21}\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2}, \\
 \Rightarrow \Sigma_{21} &= \frac{\kappa\Sigma_{11}}{\kappa + \frac{(1 - \kappa)\Sigma_{11}}{\Sigma_{11} + \sigma_\xi^2}} = \frac{\kappa\Sigma_{11}(\Sigma_{11} + \sigma_\xi^2)}{\kappa(\Sigma_{11} + \sigma_\xi^2) + (1 - \kappa)\Sigma_{11}}, \\
 &= \frac{\kappa\Sigma_{11}(\Sigma_{11} + \sigma_\xi^2)}{\kappa\sigma_\xi^2 + \Sigma_{11}} = \frac{2\kappa\Sigma_{11}(2\Sigma_{11} + 2\sigma_\xi^2)}{4\kappa\sigma_\xi^2 + 4\Sigma_{11}},
 \end{aligned}$$

³⁰ 参看 [Benchimol and Bounader \(2019\)](#) 附录 A.4 式 (85)。

$$\begin{aligned}
&= \frac{\kappa \sigma_\epsilon^2 \left(1 + \sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}}\right) \left[\sigma_\epsilon^2 \left(1 + \sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}}\right) + 2\sigma_\xi^2\right]}{4\kappa \sigma_\xi^2 + 2\sigma_\epsilon^2 \left(1 + \sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}}\right)}, \\
&= \sigma_\epsilon^2 \frac{\kappa \left(1 + \sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}}\right) \left[\left(1 + \sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}}\right) + 2 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}\right]}{4\kappa \frac{\sigma_\xi^2}{\sigma_\epsilon^2} + 2 \left(1 + \sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}}\right)}, \\
&\equiv \sigma_\epsilon^2 \frac{\kappa(1+X)[(1+X) + 2\frac{a}{b}]}{4\kappa \frac{a}{b} + 2(1+X)} = \sigma_\epsilon^2 \frac{\kappa[(1+X) + 2\frac{a}{b}]}{\frac{4\kappa \frac{a}{b}}{1+X} + 2} = \sigma_\epsilon^2 \frac{(1+X) + 2\frac{a}{b}}{4\frac{a}{b} \frac{1-X}{-4\frac{a}{b}} + \frac{2}{\kappa}} = \sigma_\epsilon^2 \frac{(X+1) + 2\frac{a}{b}}{(X-1) + \frac{2}{\kappa}}, \\
&= \sigma_\epsilon^2 \frac{\left(\sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}} + 1\right) + 2 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}}{\left(\sqrt{1 + 4 \frac{\sigma_\xi^2}{\sigma_\epsilon^2}} - 1\right) + \frac{2}{\kappa}}.
\end{aligned}$$

附录 C 福利损失函数的推导

第一步，定义偏离稳态百分比的变量，取期望，求方差，依次为：

$$\hat{y}_t \equiv y_t - y, \quad \hat{y}_{it} \equiv y_{it} - y, \quad \mathbb{E}_i \hat{y}_{it} \equiv \int_0^1 \hat{y}_{it} di, \quad \text{var}_i \hat{y}_{it} = \mathbb{E}_i \hat{y}_{it}^2 - (\mathbb{E}_i \hat{y}_{it})^2.$$

第二步，产出偏离与价格偏离的关系：

$$\begin{aligned}
Y_{it} &= \left(\frac{P_{it}}{P_t}\right)^{-\gamma_t} Y_t, \\
&\Rightarrow y_{it} = y_t - \gamma_t(p_{it} - p_t), \\
&\Rightarrow \text{var}_i(y_{it} - y_t) = \gamma_t^2 \text{var}_i(p_{it} - p_t).
\end{aligned}$$

第三步，单个厂商的预期产出：

$$\begin{aligned}
Y_t &= \left(\int_0^1 Y_{it}^{\frac{\gamma_t-1}{\gamma_t}} di\right)^{\frac{\gamma_t}{\gamma_t-1}}, \\
\Rightarrow e^{\frac{\gamma_t-1}{\gamma_t} y_t} &= \int_0^1 e^{\frac{\gamma_t-1}{\gamma_t} y_{it}} di, \\
&= \int_0^1 e^{\frac{\gamma_t-1}{\gamma_t} y} di + \int_0^1 \left[\frac{\gamma_t-1}{\gamma_t} e^{\frac{\gamma_t-1}{\gamma_t} y} \hat{y}_{it} + \frac{1}{2} \left(\frac{\gamma_t-1}{\gamma_t}\right)^2 e^{\frac{\gamma_t-1}{\gamma_t} y} \hat{y}_{it}^2 \right] di, \\
&= e^{\frac{\gamma_t-1}{\gamma_t} y} \left[1 + \frac{\gamma_t-1}{\gamma_t} \left(\mathbb{E}_i \hat{y}_{it} + \frac{1}{2} \frac{\gamma_t-1}{\gamma_t} \mathbb{E}_i \hat{y}_{it}^2 \right) \right], \\
\Rightarrow e^{\frac{\gamma_t-1}{\gamma_t} (y_t - y)} &= \left[1 + \frac{\gamma_t-1}{\gamma_t} \left(\mathbb{E}_i \hat{y}_{it} + \frac{1}{2} \frac{\gamma_t-1}{\gamma_t} \mathbb{E}_i \hat{y}_{it}^2 \right) \right], \\
\Rightarrow \frac{\gamma_t-1}{\gamma_t} \hat{y}_t &= \log \left[1 + \frac{\gamma_t-1}{\gamma_t} \left(\mathbb{E}_i \hat{y}_{it} + \frac{1}{2} \frac{\gamma_t-1}{\gamma_t} \mathbb{E}_i \hat{y}_{it}^2 \right) \right], \\
&\approx \frac{\gamma_t-1}{\gamma_t} \left(\mathbb{E}_i \hat{y}_{it} + \frac{1}{2} \frac{\gamma_t-1}{\gamma_t} \mathbb{E}_i \hat{y}_{it}^2 \right), \\
\Rightarrow \hat{y}_t &= \mathbb{E}_i \hat{y}_{it} + \frac{\gamma_t-1}{2\gamma_t} \mathbb{E}_i \hat{y}_{it}^2,
\end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E}_i \hat{y}_{it} + \frac{\gamma_t - 1}{2\gamma_t} [\text{var}_i \hat{y}_{it} + (\mathbb{E}_i \hat{y}_{it})^2], \\
 &\approx \mathbb{E}_i \hat{y}_{it} + \frac{\gamma_t - 1}{2\gamma_t} \text{var}_i \hat{y}_{it}, \\
 \Rightarrow \quad &\mathbb{E}_i \hat{y}_{it} = \hat{y}_t - \frac{\gamma_t - 1}{2\gamma_t} \text{var}_i \hat{y}_{it}.
 \end{aligned}$$

第四步，福利损失函数：

$$\begin{aligned}
 U(C_t, N_t) &= \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\nu}}{1+\nu} = \frac{Y_t^{1-\sigma}}{1-\sigma} - \int_0^1 \frac{N_{it}^{1+\nu}}{1+\nu} di, \\
 &= \frac{e^{(1-\sigma)y_t}}{1-\sigma} - \int_0^1 \frac{e^{(1+\nu)(y_{it}-a_t)}}{1+\nu} di, \\
 &\approx \frac{e^{1-\sigma}y}{1-\sigma} + e^{(1-\sigma)y}(y_t - y) + \frac{1}{2}(1-\sigma)e^{(1-\sigma)y}(y_t - y)^2 \\
 &\quad - \int_0^1 \left[\frac{e^{(1+\nu)(y-a)}}{1+\nu} + e^{(1+\nu)(y-a)}(y_{it} - y) - e^{(1+\nu)(y-a)}(a_t - a) \right. \\
 &\quad \left. + \frac{1}{2}e^{(1+\nu)(y-a)}(1+\nu)(y_{it} - y)^2 + \frac{1}{2}e^{(1+\nu)(y-a)}(a_t - a)^2 \right. \\
 &\quad \left. - e^{(1+\nu)(y-a)}(1+\nu)(y_{it} - y)(a_t - a) \right] di, \\
 &\approx e^{(1-\sigma)y} \left(\hat{y}_t + \frac{1-\sigma}{2} \hat{y}_t^2 \right) - e^{(1+\nu)(y-a)} \int_0^1 \left[\hat{y}_{it} + \frac{1+\nu}{2} \hat{y}_{it}^2 - (1+\nu) \hat{y}_{it} a_t \right] di + \text{t.i.p.}, \\
 &= e^{(1-\sigma)y} \left(\hat{y}_t + \frac{1-\sigma}{2} \hat{y}_t^2 \right) - e^{(1+\nu)(y-a)} \left[\mathbb{E}_i \hat{y}_{it} + \frac{1+\nu}{2} \mathbb{E}_i \hat{y}_{it}^2 - (1+\nu) a_t \mathbb{E}_i \hat{y}_{it} \right] + \text{t.i.p.}, \\
 &= e^{(1-\sigma)y} \left(\hat{y}_t + \frac{1-\sigma}{2} \hat{y}_t^2 \right) - e^{(1+\nu)(y-a)} \left\{ \mathbb{E}_i \hat{y}_{it} + \frac{1+\nu}{2} [\text{var}_i \hat{y}_{it} + (\mathbb{E}_i \hat{y}_{it})^2] - (1+\nu) a_t \mathbb{E}_i \hat{y}_{it} \right\} + \text{t.i.p.}, \\
 &\approx e^{(1-\sigma)y} \left(\hat{y}_t + \frac{1-\sigma}{2} \hat{y}_t^2 \right) - e^{(1+\nu)(y-a)} \left\{ \left(\hat{y}_t - \frac{\gamma_t - 1}{2\gamma_t} \text{var}_i \hat{y}_{it} \right) \right. \\
 &\quad \left. + \frac{1+\nu}{2} \left[\text{var}_i \hat{y}_{it} + \left(\hat{y}_t - \frac{\gamma_t - 1}{2\gamma_t} \text{var}_i \hat{y}_{it} \right)^2 \right] - (1+\nu) a_t \left(\hat{y}_t - \frac{\gamma_t - 1}{2\gamma_t} \text{var}_i \hat{y}_{it} \right) \right\} + \text{t.i.p.}, \\
 &= e^{(1-\sigma)y} \left(\hat{y}_t + \frac{1-\sigma}{2} \hat{y}_t^2 \right) - e^{(1+\nu)(y-a)} \left[\hat{y}_t + \frac{1+\nu}{2} \hat{y}_t^2 + \frac{\gamma_t^{-1} + \nu}{2} \text{var}_i \hat{y}_{it} - (1+\nu) a_t \hat{y}_t \right] + \text{t.i.p.}, \\
 &= e^{(1-\sigma)y} \left(\hat{y}_t + \frac{1-\sigma}{2} \hat{y}_t^2 \right) - e^{(1-\sigma)y} \left[\hat{y}_t + \frac{1+\nu}{2} \hat{y}_t^2 + \frac{\gamma_t^{-1} + \nu}{2} \text{var}_i \hat{y}_{it} - (1+\nu) a_t \hat{y}_t \right] + \text{t.i.p.}, \\
 &= -e^{(1-\sigma)y} \frac{\sigma + \nu}{2} \left(\hat{y}_t^2 - 2 \frac{1+\nu}{\sigma + \nu} a_t \hat{y}_t + \frac{\gamma_t^{-1} + \nu}{\sigma + \nu} \text{var}_i \hat{y}_{it} \right) + \text{t.i.p.}, \\
 &\approx - \left[\hat{y}_t^2 - 2 \frac{1+\nu}{\sigma + \nu} a_t \hat{y}_t + \left(\frac{1+\nu}{\sigma + \nu} a_t \right)^2 + \frac{\gamma_t^{-1} + \nu}{\sigma + \nu} \text{var}_i \hat{y}_{it} + \text{t.i.p.} \right], \\
 &= - \left[\hat{y}_t^2 - 2 \frac{1+\nu}{\sigma + \nu} a_t \hat{y}_t + (\hat{y}_t^e)^2 + \frac{\gamma_t^{-1} + \nu}{\sigma + \nu} \text{var}_i \hat{y}_{it} + \text{t.i.p.} \right], \\
 &= - \left[\hat{y}_t^2 - 2 \hat{y}_t^e \hat{y}_t + (\hat{y}_t^e)^2 + \frac{\gamma_t^{-1} + \nu}{\sigma + \nu} \text{var}_i \hat{y}_{it} + \text{t.i.p.} \right], \\
 &\approx - \left[(\hat{y}_t - \hat{y}_t^e)^2 + \frac{\gamma_t^{-1} + \nu}{\sigma + \nu} \text{var}_i (\hat{y}_{it} - \hat{y}_t) + \text{t.i.p.} \right],
 \end{aligned}$$

$$\begin{aligned} &\approx - \left[(\hat{y}_t - \tilde{y}_t^e)^2 + \frac{\gamma_t^{-1} + \psi}{\sigma + \nu} \gamma_t^2 \text{var}_i(\hat{p}_{it} - \hat{p}_t) + \text{t.i.p.} \right], \\ &\approx - [(\tilde{y}_t^e)^2 + \lambda^{\text{par}} \text{var}_i(\hat{p}_{it} - \hat{p}_t)]. \end{aligned}$$

其中, $\lambda^{\text{par}} = \frac{\gamma_t^{-1} + \psi}{\sigma + \nu} \gamma_t^2 = \frac{\gamma_t + \gamma_t^2 \psi}{\sigma + \nu}$ 。

以上推导基于稳态是有效的, 即消费与劳动的边际替代率等于劳动的边际产出: $-\frac{U_N}{U_C} = MPN$ 。当成本加成冲击导致稳态扭曲时: $-\frac{U_N}{U_C} \neq MPN$, 此时假设 $-\frac{U_N}{U_C} = MPN(1 - \phi)$, $\phi = 1 - \frac{\gamma-1}{\gamma} = \frac{1}{\gamma}$ 代表一个类似于“楔子”的参数。将此条件替换后, 得到此时的福利损失函数 (即期) 为: ³¹

$$-\frac{U(C_t, N_t) - U(C, N)}{U_C C} \approx (\tilde{y}_t^e)^2 - \frac{\tilde{y}_t^e}{\gamma} + \lambda^{\text{par}} \text{var}_i(\hat{p}_{it} - \hat{p}_t).$$

³¹ 另可参考邓燕飞等 (2021) 附录 E 的推导思路。

Information Friction, Signal Processing, and Monetary Policy

Abstract: It is a controversial topic in macroeconomics whether the theory of sticky price or sticky information is better. The sticky price model is better from the perspective of fitting endogenous inertia while the sticky information model is better from the perspective of simulating exogenous inertia. The sticky price assumes that the information used on decision is perfect, and the sticky information assumes that this information is perfect with a certain probability. The theory of rational negligence directly faces the imperfect information, and compared with signal extraction, it can also weigh the cost and benefit of information processing to achieve the optimal allocation of attention. When the ratio of probability of stickiness to non-sticky probability's square is equal to the ratio of noise variance and shock variance, sticky price can be regarded as a special case of rational inattention. This paper proves that sticky information can also be regarded as a special case of rational inattention when the probability of updating information is equal to the Kalman gain. In the standard three equation dynamic rational inattention general equilibrium framework, we also get the impulse response diagram which is similar to the inflation and other macroeconomic variables generated by the sticky information model. Therefore, rational inattention, which has both the advantages of sticky price and sticky information, can replace the two and avoid hesitation when adopting the framework of monetary policy analysis. In particular, as far as the practical monetary policy is concerned, the welfare loss under the optimal simple rule in the rational inattention framework is smaller and closer to the ideal optimal monetary policy.

Keywords: Rational Inattention; Sticky Information; Sticky Prices; Optimal Monetary Policy; Optimal Interest Rate Rules

JEL classification: E31; E32; E52; E58