**电商市场是否会促进价格水平的稳定？**

——来自纯牛奶商品的实证证据

陈歆昱 黄益平 邱晗

**目录**

[附录I 宏观数据初步检验 1](#_Toc155443625)

[附录II 所有变量展示 2](#_Toc155443626)

[附录III 约简式回归 3](#_Toc155443627)

[附录IV 参数估计方法 9](#_Toc155443628)

[附录V 估计结果 14](#_Toc155443629)

[附录VI 结构模型稳健性检验 16](#_Toc155443630)

[附录VII 反事实模拟求解过程 2](#_Toc155443631)3

[附录VIII 关于价格分散度的讨论 2](#_Toc155443632)5

#

# 附录I 宏观数据初步检验

我们构建了2001年到2019年的省份-年度面板数据集，通过固定效应回归来进初步检验通胀波动率与电商市场发展的相关性。基准回归模型设定如下：

，

其中*c*指代省份，*t*指代年度。因变量 为*c*省在*t*年12个月度CPI同比增长的标准差，作为实现的波动率度量。 为快递网点密度（个/平方公里），作为电商在各省发展程度的代理变量。 为控制变量，包括人口、GDP等。 控制省级固定效应， 控制年度固定效应， 为扰动项。 和 是回归模型的系数。

表I1报告了回归系数的估计。通胀波动率与快递网点密度呈现显著的负相关关系。在最严格的控制下，如第6列所示，快递网点密度每增加1%，通胀波动率平均下降约0.1%。

**表I1 省级面板回归结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 　 | (1) | (2) | (3) | (4) | (5) | (6) |
| 因变量：Log(年度通胀标准差) |  |  |  |  |  |  |
| 　 | 　 | 　 | 　 | 　 | 　 | 　 |
| Log(快递网点密度) | -0.0808\*\* | -0.564\*\*\* | -0.0328 | -0.106\*\*\* | -0.105\*\*\* | -0.0948\*\* |
|  | (0.0357) | (0.114) | (0.0672) | (0.0347) | (0.0347) | (0.0349) |
| Log(人口) |  |  |  | 0.655 | 0.686\* | 0.676\* |
|  |  |  |  | (0.397) | (0.385) | (0.371) |
| Log(GDP) |  |  |  |  | 0.419 | 0.323 |
|  |  |  |  |  | (0.929) | (0.948) |
| Log(住宅平均价格) |  |  |  |  |  | 0.307\*\*\* |
|  |  |  |  |  |  | (0.0947) |
|  |  |  |  |  |  |  |
| 省级固定效应 | N | Y | Y | Y | Y | Y |
| 年度固定效应 | N | N | Y | Y | Y | Y |
|  |  |  |  |  |  |  |
| 观测值 | 558 | 558 | 558 | 558 | 558 | 558 |
| R2 | 0.047 | 0.339 | 0.721 | 0.724 | 0.724 | 0.726 |

注：标准误双重聚类到省级和年度；\*\*\* p<0.01, \*\* p<0.05, \* p<0.1。

# 附录II 所有变量展示

我们在表II1中，展示了文中所用变量的名称、单位、地区范围、频率、时间跨度和数据来源。

**表II1 所有变量展示**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | 单位 | 地区范围 | 频率 | 时间跨度 | 来源 |
| CPI同比 | % | 全国/省级 | 月度 | 2001年-2019年 | 国家统计局 |
| PPI同比 | % | 全国/省级 | 月度 | 2001年-2019年 | 国家统计局 |
| 纯牛奶多包装集合线下价格 | 元/升 | 地级市 | 月度 | 2015年-2019年 | 某市场营销公司 |
| 纯牛奶单包装集合线下价格 | 元/升 | 地级市 | 月度 | 2015年-2019年 | 某市场营销公司 |
| 纯牛奶多包装集合天猫超市价格 | 元/升 | 地级市 | 月度 | 2015年-2019年 | 阿里巴巴研究院 |
| 纯牛奶多包装集合线下销售额 | 百万元 | 地级市 | 月度 | 2015年-2019年 | 某市场营销公司 |
| 纯牛奶单包装集合线下销售额 | 百万元 | 地级市 | 月度 | 2015年-2019年 | 某市场营销公司 |
| 纯牛奶多包装集合天猫超市销售额 | 百万元 | 地级市 | 月度 | 2015年-2019年 | 阿里巴巴研究院 |
| 城市等级 | - | 地级市 | - | 2019年 | 第一财经 |
| 常住人口数 | 万人 | 省级/地级市 | 年度 | 2001-2019年 | 各地统计局 |
| 面积 | 平方公里 | 省级/地级市 | - | - | 各地统计局 |
| 人均收入 | 元/年 | 省级/地级市 | 年度 | 2001-2019年 | 各地统计局 |
| 社会消费品零售总额 | 亿元 | 地级市 | 年度 | 2019年 | 各地统计局 |
| 快递服务业务量 | 亿件 | 地级市 | 年度 | 2015-2019年 | 各地邮政行业发展公报 |
| 快递服务路线长度 | 公里 | 地级市 | 年度 | 2015-2019年 | 各地邮政行业发展公报 |
| 牛奶产量 | 万吨 | 省级 | 年度 | 2015-2019年 | 各省国民经济和社会发展统计公报 |
| 人均奶类消费量 | 千克 | 省级 | 年度 | 2015-2019年 | 各省统计年鉴 |
| 零售企业利润率 | % | 省级 | 年度 | 2015-2019年 | 各省统计年鉴 |
| 经纬度 | - | 地市级 | - | - | 高德地图 |
| 劳动年龄人口占比 | % | 地市级 | 十年 | 2020年 | 第七次人口普查 |
| 平均受教育年限 | 年 | 地市级 | 十年 | 2020年 | 第七次人口普查 |
| 夜间灯光强度 | - | 地市级 | 月度 | 2015年-2019年 | NPP-VIIRS |
| 纯牛奶B2C电商市场份额 | % | 全国 | 年度 | 2017年 | 三胜咨询 |
| B2C电商市场份额 | % | 全国 | 年度 | 2015年-2019年 | 网经社 |
| 日内最高气温 | ℃ | 地市级 | 日度 | 2015年-2019年 | Wind经济数据库 |
| 快递企业数量 | 个 | 地市级 | 月度 | 2015年-2019年 | Wind经济数据库 |
| 同城快递续重收费标准 | 元/千克 | - | 年度 | 2021年 | 买购网 |
| 企业毛利率 | % | - | 年度 | 2019年 | 企业年报 |
| GDP | 亿元 | 省级 | 年度 | 2001-2019年 | 国家统计局 |
| 快递网点数量 | 个 | 省级 | 年度 | 2001-2019年 | Wind经济数据库 |
| 住宅平均价格 | 元/平方米 | 省级 | 年度 | 2001-2019年 | Wind经济数据库 |

# 附录III 约简式回归

为增强约简式回归的稳健性，本文还使用了 作为因变量的替代性指标，来衡量价格波动率。 为地级市*n*的线下商超中，商品集合*j*在某年*t*中12个月的去趋势价格的四分位距。由于计算IQR统计量用到的信息不及标准差全面，因此 衡量波动率的效率有所下降。基准回归模型中的结果基本不变，但在“市场融合”渠道的检验中，单独针对多包装的IV回归系数在10%的显著性水平上不再显著（但在15%的显著性水平上显著），联合回归的IV估计仍然显著；在“竞争效应”渠道的检验中，OLS回归的系数不显著，但IV回归仍然是高度显著的。

**表III1 基准模型回归结果（IQR衡量年度价格波动）**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 　 | (1) | (2) | 　 | (3) | (4) | 　 | (5) | (6) |
|  | 多包装 | 　 | 单包装 | 　 | 联合回归 |
| *因变量：Log（年度价格波动）* | OLS | IV | 　 | OLS | IV | 　 | OLS | IV |
|  |  |  |  |  |  |  |  |  |
| 天猫销售额占比 | -5.048\* | -14.10\*\*\* |  | 0.215 | 0.015 |  | 3.566 | 0.189 |
|  | (2.741) | (4.904) |  | (3.915) | (6.234) |  | (2.824) | (4.296) |
| 天猫销售额占比\*多包装 |  |  |  |  |  |  | -11.93\*\*\* | -14.46\*\*\* |
|  |  |  |  |  |  |  | (2.810) | (3.250) |
|  |  |  |  |  |  |  |  |  |
| 控制变量 | Y | Y |  | Y | Y |  | Y | Y |
| 年度固定效应 | Y | Y |  | Y | Y |  | Y | Y |
| 城市固定效应 | Y | Y |  | Y | Y |  |  |  |
| 商品集合\*城市固定效应 |  |  |  |  |  |  | Y | Y |
|  |  |  |  |  |  |  |  |  |
| 样本量 | 296 | 214 |  | 295 | 214 |  | 591 | 428 |
| R2 | 0.600 | 　 | 　 | 0.552 | 　 | 　 | 0.558 | 　 |
|  |

注：括号中为标准误，单独回归在城市上聚类，联合回归在商品集合-城市上聚类；\*\*\* p<0.01, \*\* p<0.05, \* p<0.1。

**表III2 “市场融合”渠道回归结果（IQR衡量年度价格波动）**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 　 | (1) | (2) | 　 | (3) | (4) | 　 | (5) | (6) |
|  | 多包装 | 　 | 单包装 | 　 | 联合回归 |
| *因变量：Log（年度价格波动）* | OLS | IV | 　 | OLS | IV | 　 | OLS | IV |
|  |  |  |  |  |  |  |  |  |
| 天猫销售额占比 | -3.075 | -8.096 |  | -0.893 | -5.355 |  | -0.613 | -6.725 |
|  | (4.170) | (6.584) |  | (6.751) | (11.55) |  | (3.933) | (7.307) |
| 天猫销售额占比\*牛奶进口省份 | -5.570 | -14.33 |  | -2.190 | 6.803 |  | -0.632 | 11.91 |
|  | (8.655) | (9.388) |  | (12.95) | (15.84) |  | (10.34) | (12.52) |
| 天猫销售额占比\*牛奶进口省份\*多包装 |  |  |  |  |  |  | -8.447 | -31.34\*\*\* |
|  |  |  |  |  |  |  | (9.294) | (10.06) |
|  |  |  |  |  |  |  |  |  |
| 控制变量 | Y | Y |  | Y | Y |  | Y | Y |
| 年度固定效应 | Y | Y |  | Y | Y |  | Y | Y |
| 城市固定效应 | Y | Y |  | Y | Y |  |  |  |
| 商品集合\*城市固定效应 |  |  |  |  |  |  | Y | Y |
|  |  |  |  |  |  |  |  |  |
| 样本量 | 296 | 214 |  | 295 | 214 |  | 591 | 428 |
| R2 | 0.600 | 　 | 　 | 0.554 | 　 | 　 | 0.562 | 　 |

注：括号中为标准误，单独回归在城市上聚类，联合回归在商品集合-城市上聚类；\*\*\* p<0.01, \*\* p<0.05,

\* p<0.1。

**表III3 “竞争效应”渠道回归结果（IQR衡量年度价格波动）**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 　 | (1) | (2) | 　 | (3) | (4) | 　 | (5) | (6) |
|  | 多包装 | 　 | 单包装 | 　 | 联合回归 |
| *因变量：Log（年度价格波动）* | OLS | IV | 　 | OLS | IV | 　 | OLS | IV |
|  |  |  |  |  |  |  |  |  |
| 天猫销售额占比 | -2.209 | -3.610 |  | 1.723 | 0.522 |  | -0.248 | -1.544 |
|  | (3.068) | (4.321) |  | (4.773) | (8.265) |  | (2.737) | (5.011) |
| 天猫销售额占比\*零售企业利润率 | -1.693 | -3.663\*\*\* |  | -0.945 | -0.545 |  | -0.485 | 0.452 |
|  | (1.200) | (0.579) |  | (1.037) | (1.097) |  | (0.873) | (0.600) |
| 天猫销售额占比\*零售企业利润率\*多包装 |  |  |  |  |  |  | -1.535 | -5.111\*\*\* |
|  |  |  |  |  |  |  | (1.290) | (0.792) |
|  |  |  |  |  |  |  |  |  |
| 控制变量 | Y | Y |  | Y | Y |  | Y | Y |
| 年度固定效应 | Y | Y |  | Y | Y |  | Y | Y |
| 城市固定效应 | Y | Y |  | Y | Y |  |  |  |
| 商品集合\*城市固定效应 |  |  |  |  |  |  | Y | Y |
|  |  |  |  |  |  |  |  |  |
| 样本量 | 296 | 214 |  | 295 | 214 |  | 591 | 428 |
| R2 | 0.607 | 　 | 　 | 0.555 | 　 | 　 | 0.564 | 　 |

注：括号中为标准误，单独回归在城市上聚类，联合回归在商品集合-城市上聚类\*\*\* p<0.01, \*\* p<0.05, \* p<0.1。

所有工具变量回归的一阶段回归结果，以及外生性和有效性检验如下所示。工具变量人均快递使用量在一阶段对内生变量的回归中，其系数几乎都是高度显著的；另一工具变量快递路线密度，其系数显著的情况不算多，原因是其与人均快递使用量具有一定程度的共线性，一起进入回归时其显著性削弱（单独回归时是显著的）。尽管如此，我们仍将该变量始终作为工具变量其一进入回归。否则的话，无法进行工具变量有效性检验（检验的前提是工具变量数量大于内生变量数量）。

我们使用Sanderson-Windmeijer multivariate F test进行工具变量相关性检验，该检验的原假设是工具变量与内生变量不相关，P值越小，越倾向于拒绝原假设，倾向于认为工具变量能通过相关性检验。该检验仅与一阶段回归有关。我们使用Hansen J statistics进行工具变量外生性检验，该检验的原假设是工具变量与扰动项不相关，P值越大，越难以拒绝原假设，倾向于认为工具变量能通过外生性检验。该检验由于考察工具变量与残差的相关性，因此会随着因变量选取改变而变化。下表中的结果显示，在所有的设定下，工具变量相关性和外生性检验均能通过。

**表III4 基准模型IV一阶段回归及IV相关性、外生性检验结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 　 | (1) | 　 | 　 | (2) | (3) |
|  | 分别回归 | 　 | 　 | 联合回归 |  |
|  | (A) | 　 | 　 | (A) | (B) |
|  |  |  |  |  |  |  |
| 人均快递使用量 | 4.558\*\*\* |  |  | 4.558\*\*\* | -1.744\*\*\* |
|  | (1.283) |  |  | (1.125) | (0.574) |
| 快递路线密度 | -0.064 |  |  | -0.064 | 0.026 |
|  | (0.093) |  |  | (0.0001) | (0.031) |
| 人均快递使用量\*多包装 |  |  |  |  | 6.03e-19 | 8.046\*\*\* |
|  |  |  |  |  | (1.351) | (1.083) |
| 快递路线密度\*多包装 |  |  |  |  | 1.11e-19 | -0.117 |
|  |  |  |  |  | (0.131) | (0.118) |
|  |  |  |  |  |  |  |
| 控制变量 | Y |  |  | Y | Y |
| 年度固定效应 | Y |  |  | Y | Y |
| 城市固定效应 | Y |  |  |  |  |
| 商品集合\*城市固定效应 |  |  |  |  | Y | Y |
| 样本量 | 214 |  |  | 428 | 428 |
|  |  |  |  |  |  |  |
| Sanderson-Windmeijer multivariate F test（相关性检验） | 8.67\*\*\* |  |  | 8.82\*\*\* | 16.95\*\*\* |
| Hansen J statistics（外生性检验） | 多包装 | 单包装 |  |  | 所有包装 |
| 标准差衡量价格波动率 | 0.388 | 0.893 |  |  | 1.395 |
| IQR衡量价格波动率 | 0.964 | 1.030 |  |  | 1.906 |
|  |

注：括号中为标准误，单独回归在城市上聚类，联合回归在商品集合-城市上聚类；\*\*\* p<0.01, \*\* p<0.05, \* p<0.1；工具变量(A)代表天猫销售额占比，工具变量(B)代表天猫销售额占比\*多包装，二者具有潜在内生性；相关性检验的原假设是工具变量与内生变量不相关；外生性检验的原假设是工具变量与扰动项不相关；检验统计量的P值越小越倾向于拒绝原假设。

**表III5 “市场融合”渠道IV一阶段回归及IV相关性、外生性检验结果**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 　 | (1) | (2) | 　 | 　 | (3) | (4) | (5) |
|  | 分别回归 | 　 | 　 | 联合回归 |
|  | (A) | (B) | 　 | 　 | (A) | (B) | (C) |
|  |  |  |  |  |  |  |  |
| 人均快递使用量 | 1.687 | -4.687\*\*\* |  |  | 1.687 | -4.687\*\*\* | -2.343\*\*\* |
|  | (1.922) | (1.387) |  |  | (1.343) | (0.970) | (0.696) |
| 快递路线密度 | -0.208 | 0.271\*\*\* |  |  | -0.208 | 0.271\*\*\* | 0.136\*\* |
|  | (0.233) | (0.107) |  |  | (0.163) | (0.075) | (0.064) |
| 人均快递使用量\*牛奶进口省份 | 7.687 | 18.54\*\*\* |  |  | 7.687\* | 18.54\*\*\* | 5.090\*\*\* |
|  | (5.686) | (3.870) |  |  | (4.387) | (2.870) | (1.818) |
| 快递路线密度\*牛奶进口省份 | 0.423 | -0.864\*\*\* |  |  | 0.423 | -0.864\*\*\* | -0.365\* |
|  | (0.694) | (0.312) |  |  | (0.516) | (0.228) | (0.187) |
| 人均快递使用量\*牛奶进口省份\*多包装 |  |  |  |  | -1.37e-15 | 7.96e-15 | 8.360\*\*\* |
|  |  |  |  |  | (3.715) | (1.279) | (1.066) |
| 快递路线密度\*牛奶进口省份\*多包装 |  |  |  |  | 1.40e-16 | -4.31e-16 | -0.134 |
|  |  |  |  |  | (0.350) | (0.132) | (0.119) |
|  |  |  |  |  |  |  |  |
| 控制变量 | Y | Y |  |  | Y | Y | Y |
| 年度固定效应 | Y | Y |  |  | Y | Y | Y |
| 城市固定效应 | Y | Y |  |  |  |  |  |
| 商品集合\*城市固定效应 |  |  |  |  | Y | Y | Y |
| 样本量 | 214 | 214 |  |  | 428 | 428 | 428 |
|  |  |  |  |  |  |  |  |
| Sanderson-Windmeijer multivariate F test（相关性检验） | 6.73\*\*\* | 8.07\*\*\* |  |  | 12.38\*\*\* | 10.53\*\*\* | 24.73\*\*\* |
| Hansen J statistics（外生性检验） | 多包装 | 单包装 |  |  | 所有包装 |
| 标准差衡量价格波动率 | 1.371 | 2.451 |  |  | 0.958 |
| IQR衡量价格波动率 | 1.509 | 3.404 |  |  | 2.280 |
|  |

注：括号中为标准误，单独回归在城市上聚类，联合回归在商品集合-城市上聚类；\*\*\* p<0.01, \*\* p<0.05, \* p<0.1；工具变量(A)代表天猫销售额占比，工具变量(B)代表天猫销售额占比\*牛奶进口省份，工具变量(C)代表天猫销售额占比\*牛奶进口省份\*多包装，三者具有潜在内生性；相关性检验的原假设是工具变量与内生变量不相关；外生性检验的原假设是工具变量与扰动项不相关；检验统计量的P值越小越倾向于拒绝原假设。

**表III6 “竞争效应” 渠道IV一阶段回归及IV相关性、外生性检验结果**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 　 | (1) | (2) | 　 | 　 | (3) | (4) | (5) |
|  | 分别回归 | 　 | 　 | 联合回归 |
|  | (A) | (B) | 　 | 　 | (A) | (B) | (C) |
|  |  |  |  |  |  |  |  |
| 人均快递使用量 | 4.409\*\*\* | -2.212 |  |  | 4.409\*\*\* | -2.212 | -1.106 |
|  | (1.506) | (2.652) |  |  | (1.053) | (1.854) | (1.331) |
| 快递路线密度 | -0.002 | -0.227 |  |  | -0.002 | -0.227 | -0.114 |
|  | (0.159) | (0.324) |  |  | (0.111) | (0.226) | (0.181) |
| 人均快递使用量\*零售企业利润率 | 0.040 | 4.565\*\*\* |  |  | -0.040 | 4.565\*\*\* | -0.084 |
|  | (0.209) | (0.318) |  |  | (0.183) | (0.293) | (0.185) |
| 快递路线密度\*零售企业利润率 | -0.036 | 0.127 |  |  | -0.036 | 0.127 | 0.055 |
|  | (0.063) | (0.154) |  |  | (0.049) | (0.111) | (0.087) |
| 人均快递使用量\*零售企业利润率\*多包装 |  |  |  |  | 3.17e-16 | 5.06e-19 | 4.734\*\*\* |
|  |  |  |  |  | (0.222) | (0.382) | (0.343) |
| 快递路线密度\*零售企业利润率\*多包装 |  |  |  |  | -1.09e-16 | -5.09e-19 | 0.017 |
|  |  |  |  |  | (0.044) | (0.059) | (0.053) |
|  |  |  |  |  |  |  |  |
| 控制变量 | Y | Y |  |  | Y | Y | Y |
| 年度固定效应 | Y | Y |  |  | Y | Y | Y |
| 城市固定效应 | Y | Y |  |  |  |  |  |
| 商品集合\*城市固定效应 |  |  |  |  | Y | Y | Y |
| 样本量 | 214 | 214 |  |  | 428 | 428 | 428 |
|  |  |  |  |  |  |  |  |
| Sanderson-Windmeijer multivariate F test（相关性检验） | 6.40\*\*\* | 20.90\*\*\* |  |  | 9.82\*\*\* | 31.94\*\*\* | 72.52\*\*\* |
| Hansen J statistics（外生性检验） | 多包装 | 单包装 |  |  | 所有包装 |
| 标准差衡量价格波动率 | 2.526 | 3.771 |  |  | 3.621 |
| IQR衡量价格波动率 | 3.806 | 1.431 |  |  | 2.676 |

注：括号中为标准误，单独回归在城市上聚类，联合回归在商品集合-城市上聚类；\*\*\* p<0.01, \*\* p<0.05, \* p<0.1；工具变量(A)代表天猫销售额占比，工具变量(B)代表天猫销售额占比\*零售企业利润率，工具变量(C)代表天猫销售额占比\*零售企业利润率\*多包装，三者具有潜在内生性；相关性检验的原假设是工具变量与内生变量不相关；外生性检验的原假设是工具变量与扰动项不相关；检验统计量的P值越小越倾向于拒绝原假设。

# 附录IV 参数估计方法

在本小节中，我们将介绍需求参数和边际成本是如何估计的。由于价格和数量同时由需求方程和供给方程决定，所以OLS估计将是非一致的——以误差项形式表现的需求冲击将不可避免地通过一阶条件影响零售商定价，从而导致使得误差项与价格存在强相关性。为解决内生性问题，我们寻找价格的工具变量来估计需求方程的参数。

第一组工具变量是成本冲击。成本冲击通过一阶条件影响价格，但与需求冲击无关。我们使用某城市-月份中，日内最高气温超过30℃的天数作为工具变量。气温是影响牛奶产量的重要因素，具体说明见前文供给侧章节。

第二组工具变量是同期某商品集合在其他地区的价格，也被称为豪斯曼工具变量（Hausman IV）。其背后的思想是，零售商的进货成本、或是促销活动在不同城市间常常统一（至少存在强相关性），从而各城市的价格也是相关的，满足工具变量的相关性条件。但是有些因素，例如全国性的广告，会同时影响某一产品在不同区域的需求，从而破坏外生性条件，削弱豪斯曼工具变量的有效性。为了缓解这个问题，我们充分利用面板数据的特性，选择8个地理距离最远城市的价格，作为某特定城市的工具变量。我们认为，城市之间的地理距离越远，需求冲击的相关性越小。由于线上价格不随城市变化，因此豪斯曼工具变量只适用于线下价格。

下面介绍具体的估计方法。首先考虑第二阶段需求方程。我们遵循Berry（1994）的方法，首先通过可观测的市场份额倒推平均效用：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅳ1) |

这使得原本非线性的方程可以被线性估计：

|  |  |  |
| --- | --- | --- |
|  | *,* | (Ⅳ2) |

其中加入时间趋势项，是为了更好地识别出价格相关的参数，这些参数直接关系到弹性计算。为了直观说明，下图展示了多包装商品集合，在全国平均计算的对数零售商份额，及其价格的走势：



**图IV1 对数零售商份额与对数价格（多包装）**

图IV1清晰地显示，零售商份额与价格，在短期内呈现出明显的负相关性：价格升高，份额降低，反之亦然。但如果直接将对数份额对对数价格回归，可能得不到准确的系数估计。特别是对于线下商超来说，对数份额与对数价格在长期中存在共同的下行趋势，这会影响到回归结果。因此有必要加入时间趋势项，来吸收掉长期的下行趋势。去趋势后的多包装的对数零售商份额与价格走势如图IV2所示：



**图IV2 去趋势的对数零售商份额与对数价格（多包装）**

显然去趋势以后，对数零售商份额与对数价格的协方差显著为负，这将帮助我们更准确地识别价格相关参数。在具体的估计中，即天猫超市以外的其他电商的市场份额，在我们的数据集中观察不到的。由于我们不希望对的假设，对参数估计产生较大影响，故我们假设是线性的：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅳ3) |

这样能完全被(2)式的固定效应项和时间趋势项系数吸收，从而不影响系数估计。但正如下文将要说明的，由于的取值不可避免地会影响到第一阶段的需求方程的系数估计，我们仍需要对和做出假设：

|  |  |  |
| --- | --- | --- |
|  | ， | (Ⅳ4) |
|  | . | (Ⅳ5) |

该假设的经济含义是，这些电商在每个城市的零售额（revenue），都与天猫超市以相同的速度线性增长，并且在样本开始（t=0）和结束时期（t=T），都占据30%的网络零售市场份额。基于该设定，我们可以计算出其他电商在每个城市-月度的多包装零售额，从而进一步计算出各零售商的市场份额：

|  |  |  |
| --- | --- | --- |
|  |  *.* | (Ⅳ6) |

完成(Ⅳ6)式的计算后，我们将所有变量代入线性的(Ⅳ2)式，进行工具变量估计，即可得到第二阶段所有参数的估计值。

接下来进入第一阶段需求方程的参数估计。首先是数据的准备，我们对不同类型的零售商进行加总，包括线下、天猫超市和其他电商，得到多包装商品集合的总销售额：

|  |  |  |
| --- | --- | --- |
|  |  *.* | (Ⅳ7) |

然后再计算多包装、单包装各自的占比，从而得到第一阶段各包装形式的支出份额：

|  |  |  |
| --- | --- | --- |
|  |  *.* | (Ⅳ8) |

总体价格指数 是其他参数 的函数，导致参数非线性的进入方程，大大增加了估计难度。但Deaton and Muellbauer（1980）的论文显示，在估计过程中使用斯通价格指数来近似的效果很好，对估计结果几乎没有影响，因此我们可以计算：

|  |  |  |
| --- | --- | --- |
|  | . | (Ⅳ9) |

最后再加总得到总支出。至此我们完成了所有的数据准备工作，可以进行线性的估计。考虑到许多固定效应项和一系列参数约束，使用一般矩方法（GMM）会极其繁琐，因此我们最终采用两步骤有约束的最小二乘法（CLS）来估计参数。具体地说，第一步将价格变量对工具变量进行回归，得到预测的价格变量；第二步的约束最小二乘法回归中，使用预测的价格变量替代原始价格。

在第一阶段对的假设，的确关系到第二阶段的数据计算，从而影响第二阶段需求方程的参数估计。但由于在我们的样本中，电商销售额占比较低，因此对总体数据的影响很小。为增强稳健性，我们也考虑放松对天猫超市始终占据30%的网络零售份额的假设，允许这一份额在时间维度上变化。这样尽管破坏了线性结构，会对第二阶段的参数估计产生实质性影响，但可能会帮助对第一阶段的需求方程进行更准确的估计。对应的估计结果见附录Ⅵ。

我们使用自助法（bootstrap）计算标准误。为了不破坏产品维度和时间维度的结构，我们有放回地对城市进行抽样，随机抽取出有重复的N个城市（例：北京、北京、上海……），构建出新的面板数据样本。通常认为200次的抽样已经足够（Chaudhuri et al., 2006），但为了保险起见，我们对原始数据集进行了300次允许重复的重新抽样，计算各个参数在这300个样本下的估计值，这300个估计值的标准差即为自助法得到的标准误。

得到参数估计后，我们进入对需求价格弹性的计算。我们有以下偏导数：

再根据链式法则，可以计算得到需求价格弹性。

最后进入供给侧估计。给定参数化的边际成本方程和矩条件，可以使用矩方法进行参数估计。

# 附录V 估计结果

表V1报告了结构模型中参数的估计值。可以看到，劳动年龄人口占比越低、人均收入越高的城市，其对价格的敏感度越低；气温越高，牛奶生产成本越高；当地快递企业密度越大，天猫超市的边际成本越低。这些都与经济学认知相符。

**表V1 结构模型参数估计值**

|  |
| --- |
| 一阶段需求 |
| 商品集合 | 对数价格 | 预算 |
| 多包装 | 单包装 |
| 多包装 | -0.431\*\*\* | 0.431\*\*\* | 0.020\*\*\* |
|  | (0.144) | (0.144) | (0.008) |
| 单包装 | 0.431\*\*\* | -0.431\*\*\* | -0.020\*\*\* |
| 　 | (0.144) | (0.144) | (0.008) |
|  |  |  |  |
| 二阶段需求 |
| 因变量：平均效用 |
| Log(价格) | -5.034\*\*\* |
|  |  |  | (0.271) |
| Log(价格)\*劳动年龄人口占比 | -0.092\*\*\* |
|  |  |  | (0.028） |
| Log(价格)\*平均受教育年限 | 0.104 |
|  |  |  | (0.180) |
| Log(价格)\*Log(人均年收入) | 2.723\*\*\* |
|  |  |  | (0.756) |
| Log(夜间灯光强度)\*1(线下零售) | -0.231\*\*\* |
|  |  |  | (0.029) |
| Log(夜间灯光强度)\*1(天猫超市) | 0.143\*\*\* |
| 　 |  | 　 | (0.029) |
| 零售商\*城市固定效应 | Y |
|  |  |  |  |
| 供给侧方程 |
| 因变量：边际成本 |  |
| 日内最高温度超过30℃的天数 | 0.015\*\*\* |
|  |  |  | (0.005) |
| 快递企业密度\*1(天猫超市) | -4.775\*\*\* |
|  |  |  | (0.180) |
| 商品集合\*城市固定效应 | Y |

注：括号里的标准误由自助法获得\*\*\* p<0.01, \*\* p<0.05, \* p<0.1。

给定需求方程的参数后，可以进一步分析隐含需求价格弹性。由于隐含需求价格弹性是价格和销售额占比的函数，随时间和地区不同，因此我们在表V2报告在平均价格和平均销售额上估计的隐含需求价格弹性。表V2的对角线单元格展示的是自身价格弹性，它们在数值上都是负的。与线下相比，线上的需求价格弹性似乎特别高，这表明消费者可能对线上价格相当敏感。其余单元格展示的是交叉需求价格弹性，所有的交叉需求价格弹性都是正的，且具有经济意义，这反映了线上-线下零售间的高度替代性。

**表V2 隐含需求价格弹性估计**

|  |  |
| --- | --- |
| 需求弹性 | 商品集合价格变动 |
| 多包装（天猫） | 多包装（线下） | 单包装（线下） |
| 多包装（天猫） | -8.006 | 5.539 | 0.569 |
|  | (0.513) | (0.479) | (0.198) |
| 多包装（线下） | 0.205 | -2.672 | 0.569 |
|  | (0.029) | (0.208) | (0.198) |
| 单包装（线下） | 0.141 | 1.648 | -2.709 |
| 　 | (0.043) | (0.500) | (0.546) |
| 注：隐含需求价格弹性在平均价格和平均销售额上估计；括号里的标准误由自助法获得。 |

模型中的边际成本同样随地区和时间而不同，为了与表V2中报告的隐含需求价格弹性保持一致，我们在表V3中报告在平均价格和平均销售额上“隐含的”边际成本和毛利率。表V3的估计结果显示，天猫超市的边际成本比线下零售更高。这或许主要是因为天猫超市在很多情况下，承担了“最后一公里”的运费。二者之差为2.5元/升，基本也与公开数据吻合[[1]](#footnote-0)。天猫超市的加价（markup）和毛利率显著比线下零售低，一方面是因为网络零售的长尾效应，常常薄利多销；另一方面也是由于线下实体店的仓储、店租等固定成本较高，零售必须要有较高的毛利率，才能覆盖这些固定成本（典型例子：眼镜店）。我们也对真实世界可比数据进行了比照，也是较为接近的[[2]](#footnote-1)。

**表V3 边际成本、加价及毛利率估计**

|  |  |  |  |
| --- | --- | --- | --- |
| 　 | 边际成本 | 加价（Markup） | 毛利率 |
| 多包装（天猫） | 14.46 | 2.05 | 12.5% |
|  | （0.15） | （0.13） | （0.01） |
| 多包装（线下） | 11.91 | 7.12 | 37.4% |
|  | （0.59） | （0.60） | （0.03） |
| 单包装（线下） | 12.86 | 7.53 | 36.9% |
| 　 | （2.22） | （2.22） | （0.11） |
| 注：边际成本、加价和毛利率在平均价格和平均销售额上估计；括号里的标准误由自助法获得。 |

# 附录VI 结构模型稳健性检验

**时变（time-varying）天猫超市份额的估计**

我们能获取的天猫超市在纯牛奶品类上的网络零售市场份额数据，仅有2017年一年，而能获取的天猫超市在全品类的网络零售市场份额数据，可以自2015年跨度至2019年。我们假定存在平行趋势：

|  |  |  |
| --- | --- | --- |
|  | ， | (Ⅵ1) |

其中代表天猫超市在全品类的网络零售市场份额，2015年至2019年分别为57.4%，57.7%，52.7%，53.5%，50.1%[[3]](#footnote-2)，据此我们计算时变的天猫超市在纯牛奶品类上的网络零售市场份额，为33.7%，33.8%，30.9%，31.4%，29.4%。然后根据这些份额比例，我们对每一年，分别计算出其他电商在每个城市-月度的多包装零售额，并进一步计算出其他变量的数据。准备完数据后，我们对参数、需求价格弹性和边际成本进行估计，估计结果与基准结构模型设定下的结果都是高度相似的。反事实模拟依赖于参数、需求价格弹性和边际成本的估计，无外部选项下的结果与基准结构模型也相仿，为避免冗长，不再额外报告。

**表VI1 结构模型参数估计值（时变天猫超市份额）**

|  |
| --- |
| 一阶段需求 |
| 商品集合 | 对数价格 | 预算 |
| 多包装 | 单包装 |
| 多包装 | -0.499\*\*\* | 0.499\*\*\* | 0.019\*\* |
|  | (0.154) | (0.154) | (0.009) |
| 单包装 | 0.499\*\*\* | -0.499\*\*\* | -0.019\*\* |
| 　 | (0.154) | (0.154) | (0.009) |
|  |  |  |  |
| 二阶段需求 |
| 因变量：平均效用 |
| Log(价格) | -6.772\*\*\* |
|  |  |  | (0.554) |
| Log(价格)\*劳动年龄人口占比 | -0.122 |
|  |  |  | (0.088） |
| Log(价格)\*平均受教育年限 | 0.250 |
|  |  |  | (0.530) |
| Log(价格)\*Log(人均年收入) | 0.819\*\* |
|  |  |  | (2.526) |
| Log(夜间灯光强度)\*1(线下零售) | -0.208\*\*\* |
|  |  |  | (0.047) |
| Log(夜间灯光强度)\*1(天猫超市) | 0.157\*\* |
|  |  |  | (0.071) |
|  |  |  |  |
| 供给侧方程 |
| 因变量：边际成本 |  |
| 日内最高温度超过30℃的天数 | 0.012\*\*\* |
|  |  |  | (0.004) |
| 快递企业密度\*1(天猫超市) | -4.739\*\*\* |
|  |  |  | (0.182) |

注：括号里的标准误由自助法获得；\*\*\* p<0.01, \*\* p<0.05, \* p<0.1。

**表VI2 隐含需求价格弹性估计（时变天猫超市份额）**

|  |  |
| --- | --- |
| 需求弹性 | 商品集合价格变动 |
| 多包装（天猫） | 多包装（线下） | 单包装（线下） |
| 多包装（天猫） | -7.579 | 5.275 | 0.663 |
|  | (0.536) | (0.501) | (0.213) |
| 多包装（线下） | 0.192 | -2.497 | 0.663 |
|  | (0.030) | (0.219) | (0.213) |
| 单包装（线下） | 0.154 | 1.861 | -2.941 |
| 　 | (0.044) | (0.524) | (0.574) |
| 注：隐含需求价格弹性在平均价格和平均销售额上估计；括号里的标准误由自助法获得。 |

**表VI3 边际成本、加价与毛利率估计（时变天猫超市份额）**

|  |  |  |  |
| --- | --- | --- | --- |
| 　 | 边际成本 | 加价（Markup） | 毛利率 |
| 多包装（天猫） | 14.34 | 2.18 | 13.2% |
|  | （0.17） | （0.15） | （0.01） |
| 多包装（线下） | 11.41 | 7.62 | 40.1% |
|  | （0.75） | （0.76） | （0.04） |
| 单包装（线下） | 13.45 | 6.93 | 34.0% |
| 　 | （2.27） | （2.27） | （0.11） |
| 注：边际成本、成本加成率和毛利率在平均价格和平均销售额上估计；括号里的标准误由自助法获得。 |

**无外部选项（天猫超市与其他电商合并）的估计**

考虑到电商之间纯牛奶商品的价格差异可能比较小，也可以使用天猫超市的价格代表电商整体的价格，并把其他电商的销售额与天猫超市合并，将电商渠道作为一个整体进行分析，从而在二阶段中取消了外部选项（两个选项时不存在IIA问题）。这样做的优点在于避免Logit模型通常被人诟病的IIA问题，但缺点在于忽略了电商间的竞争关系，且无法进行关闭天猫超市的反事实分析。

我们也汇报在这一设定下，参数、需求价格弹性和边际成本的估计结果，这些与基准结构模型设定下的结果都是高度相似的。尤其是一阶段需求的参数不受影响，针对单包装的需求价格弹性和边际成本也不变。反事实模拟依赖于参数、需求价格弹性和边际成本的估计，无外部选项下的结果与基准结构模型也相仿，为避免冗长，不再额外报告。

**表VI4 结构模型参数估计值（无外部选项）**

|  |
| --- |
| 一阶段需求 |
| 商品集合 | 对数价格 | 预算 |
| 多包装 | 单包装 |
| 多包装 | -0.431\*\*\* | 0.431\*\*\* | 0.020\*\*\* |
|  | (0.144) | (0.144) | (0.008) |
| 单包装 | 0.431\*\*\* | -0.431\*\*\* | -0.020\*\*\* |
| 　 | (0.144) | (0.144) | (0.008) |
|  |  |  |  |
| 二阶段需求 |
| 因变量：平均效用 |
| Log(价格) | -5.273\*\*\* |
|  |  |  | (0.237) |
| Log(价格)\*劳动年龄人口占比 | -0.023 |
|  |  |  | (0.032） |
| Log(价格)\*平均受教育年限 | -0.394\* |
|  |  |  | (0.205) |
| Log(价格)\*Log(人均年收入) | 2.249\*\* |
|  |  |  | (0.879) |
| Log(夜间灯光强度)\*1(线下零售) | -0.369\*\*\* |
|  |  |  | (0.035) |
|  |  |  |  |
| 供给侧方程 |
| 因变量：边际成本 |  |
| 日内最高温度超过30℃的天数 | 0.013\*\*\* |
|  |  |  | (0.004) |
| 快递企业密度\*1(天猫超市) | -4.994\*\*\* |
|  |  |  | (0.221) |

注：括号里的标准误由自助法获得；\*\*\* p<0.01, \*\* p<0.05, \* p<0.1。

**表VI5 隐含需求价格弹性估计（无外部选项）**

|  |  |
| --- | --- |
| 需求弹性 | 商品集合价格变动 |
| 多包装（天猫） | 多包装（线下） | 单包装（线下） |
| 多包装（天猫） | -5.589 | 3.993 | 0.569 |
|  | (0.354) | (0.370) | (0.198) |
| 多包装（线下） | 0.668 | -2.263 | 0.569 |
|  | (0.100) | (0.164) | (0.198) |
| 单包装（线下） | 0.453 | 1.335 | -2.709 |
| 　 | (0.136) | (0.409) | (0.546) |
| 注：隐含需求价格弹性在平均价格和平均销售额上估计；括号里的标准误由自助法获得。 |

**表VI6 边际成本、加价与毛利率估计（无外部选项）**

|  |  |  |  |
| --- | --- | --- | --- |
| 　 | 边际成本 | 加价（Markup） | 毛利率 |
| 多包装（天猫） | 13.56 | 2.96 | 17.9% |
|  | （0.20） | （0.19） | （0.01） |
| 多包装（线下） | 10.62 | 8.40 | 44.2% |
|  | （0.65） | （0.66） | （0.03） |
| 单包装（线下） | 12.86 | 7.53 | 36.9% |
| 　 | （2.22） | （2.22） | （0.11） |
| 注：边际成本、成本加成率和毛利率在平均价格和平均销售额上估计；括号里的标准误由自助法获得。 |

**不忽略交叉需求价格弹性的估计**

假设线下零售商一共售卖K件商品。我们对于线下零售商的一阶条件进行变形：

|  |  |  |
| --- | --- | --- |
|  |  | （Ⅵ2) |

记线下零售商对商品集合j的零售额为，并记，我们有：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅵ3) |

将(Ⅵ3)式写成矩阵的形式：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅵ4) |

其中，，是K\*K的交叉需求价格弹性矩阵。对(Ⅵ4)式进行求解，从而得到矩阵形式的矩条件：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅵ5) |

由于我们无法获得线下零售商所有商品信息，只能得到多包装和单包装牛奶的价格和销量数据，因此(Ⅵ5)式只有前两行能被我们利用，且的求逆存在困难。一种处理方法如前文所述，忽略所有的交叉需求价格弹性，使得成为对角矩阵，方便求逆，从而简化得到显式一阶条件。另一种处理方法是对的子矩阵进行精确求逆，其中是多包装、单包装的完整交叉需求价格弹性矩阵，然后再对(Ⅵ5)式的前两行进行运算。实际上在相当一般的情况下，前一种方法对矩阵求逆的精确度更高。证明如下：

假设所有的自身需求价格弹性，都是交叉价格需求弹性的倍。从而标准化后的具有如下的形式：

容易验证， 的逆矩阵具有相同形式的结构：

其中对角线元素，非对角元素。即使是小型的零售杂货店，售卖成百上千种商品也很正常，K的数量通常很大，因此，。

在假设对角矩阵求逆时，，误差为；，误差为0；在对2\*2的子矩阵精确求逆时， ，误差为；，误差为。第二种方法求逆的误差是严格大于第一种的，尤其在d比较接近1的时候。

另一方面，在结构模型的框架下，即使由于夸大了市场力量而低估了边际成本，其影响也会在再次求解均衡价格时被部分抵消。为了增强稳健性，我们仍使用第二种精确求逆的方法得到矩条件，并得到边际成本、加价和毛利率的估计如表VI7所示。显然这些估计值与真实世界数据之间存在较大差距。

**表VI7 边际成本、加价及毛利率估计（不忽略交叉需求价格弹性）**

|  |  |  |  |
| --- | --- | --- | --- |
| 　 | 边际成本 | 加价（Markup） | 毛利率 |
| 多包装（天猫） | 14.46 | 2.06 | 12.5% |
|  | （0.17） | （0.15） | （0.01） |
| 多包装（线下） | 9.26 | 9.76 | 51.3% |
|  | （0.52） | （0.56） | （0.03） |
| 单包装（线下） | 5.95 | 14.43 | 70.8% |
| 　 | （1.15） | （1.15） | （0.06） |
| 注：边际成本、加价和毛利率在平均价格和平均销售额上估计；括号里的标准误由自助法获得。 |

我们接着在表VI7的基础上，进行反事实模拟。在天猫超市关闭的场景下，反事实的价格波动率如表VI8所示。波动率的上升幅度有了一定程度的扩大，但仍处在合理范围内，不改变本文的基本结论。

**表VI8 天猫超市关闭的反事实价格波动率**

|  |
| --- |
| 反事实场景：天猫超市关闭 |
|  | 城市内价格波动率均值 |  | 全国加总价格波动率 |
| 　 | 反事实(A) | 原始数据(B) | Log(A)-Log(B) | 　 | 反事实(A) | 原始数据(B) | Log(A)-Log(B) |
| 多包装（天猫） | N.A. | 1.118 | N.A. |  | N.A. | 1.093 | N.A. |
|  |  | (0.016) |  |  |  | (0.037) |  |
| 多包装（线下） | 0.891 | 0.786 | 12.6% |  | 0.527 | 0.463 | 12.6% |
|  | (0.036) | (0.031) | (0.019) |  | (0.041) | (0.037) | (0.025) |
| 单包装（线下） | 1.036 | 0.958 | 7.8% |  | 0.351 | 0.274 | 24.7% |
|  | (0.060) | (0.057) | (0.014) |  | (0.051) | (0.051) | (0.080) |
| 总体 | 0.843 | 0.721 | 15.8% |  | 0.526 | 0.451 | 15.5% |
| 　 | (0.035) | (0.030) | (0.020) | 　 | (0.042) | (0.037) | (0.028) |
| 注：括号里的标准误由自助法获得。 |  |  |  |  |  |

在电商市场消失的反事实场景下，由于数学性质的巧合，反事实的矩条件(Ⅵ3)式刚好无法成立。证明如下：

对(Ⅵ3)式两边同除以线下零售商多包装和单包装的总零售额。由于在电商市场消失的反事实场景下，只存在线下零售的单一渠道，因此除法运算以后，零售额实际上将变成市场份额。此时我们将有：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅵ6) |

这里。类似地，矩阵形式的矩条件将变成：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅵ7) |

巧合的是，如果不假设对角矩阵，而对(Ⅵ7)式的进行精确求逆，我们刚好有，从而意味着。证明如下：

令，。的表达式如下：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅵ8) |

容易验证，，亦即。故而在电商市场消失的反事实场景下，不忽略交叉需求价格弹性的矩条件发生退化（deteriorate），反事实价格序列无法求解。但在仅天猫超市关闭的场景下（此时(Ⅵ7)式不成立），或者忽略交叉需求价格弹性的简化一阶条件下（此时(Ⅵ8)式发生改变），反事实价格序列仍然是可解的。

# 附录VII 反事实模拟求解过程

具体的过程如下：

我们求解每个反事实情景下的新均衡价格序列，来评估电商市场的影响。核心步骤是求解反事实价格向量的不动点。为展示清晰，以下所有公式均写成向量形式。

首先考虑关闭天猫的反事实场景：

在一阶段需求方程中，各商品集合的销售额占比可以写成以向量形式表示的，价格、总预算和扰动项的函数：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅶ1) |

如果关闭天猫超市，零售渠道的市场份额函数发生变化，但同样可以写成价格、外生变量和扰动项的函数：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅶ2) |

一阶段的价格向量是二阶段价格向量以销售额为权重的加权平均。而自身需求价格弹性是价格向量和销售额占比、市场份额占比的函数：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅶ3) |

不存在网上零售时，线下零售商的一阶条件，写成逆弹性法则的形式有：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅶ4) |

结合式(Ⅶ3)、式(Ⅶ4)，我们有：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅶ5) |

在参数、总预算、需求冲击、边际成本[[4]](#footnote-3)都已得到估计的情况下， 实际上是(Ⅶ5)式的不动点。在每一个*n*和*t*上求解方程的数值解[[5]](#footnote-4)，即可得到反事实价格。许多科学计算软件（如Matlab、Octave）内置的数值求解工具都可以完成这一工作。

关闭整体网络零售市场的反事实价格求解过程也是类似的。此时由于只有实体商超，第二阶段的需求消失。此时需求价格弹性函数简化为：

|  |  |  |
| --- | --- | --- |
|  |  | (Ⅶ6) |

我们可以将自身需求价格弹性写成以向量形式表示的函数：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅶ7) |

再结合(Ⅶ4)式，我们得到关于反事实价格向量的不动点形式：

|  |  |  |
| --- | --- | --- |
|  | *.* | (Ⅶ8) |

此时再用数值求解工具即可。

在得到反事实价格序列后，我们再进一步考察价格波动率，详见正文。

# 附录VIII 关于价格分散度的讨论

价格趋同是市场整体化程度提高的重要特征和证据。利用现有的宏观数据，我们首先检验CPI通胀趋同，是否在电商市场蓬勃发展的过去几年出现。我们基于省级CPI月度同比增长数据，计算其省间IQR（inter quartile range），以衡量通胀在各省的分散度[[6]](#footnote-5)（dispersion）。图VIII1绘制了该指标随时间变化的趋势。通胀在省间的分散度不断降低，尽管2020年初受到新冠疫情的影响有所反弹，但总体上2012年以后通胀的分散度仍显著低于之前。这一证据表明省级通胀正在趋同。

**图VIII1 省级通胀趋同（资料来源：国家统计局，作者计算）**

我们也对微观的纯牛奶数据进行考察。首先我们对价格进行方差分析（ANOVA），以考察总体价格方差中，多大部分是由时间或地区差异贡献的。结果如表VIII1所示，其中最后两列分别显示地区、时间维度对总体方差的贡献比例。对于线下价格，地区差异解释了单包装方差的54%，这一比例高于多包装的47%，这说明线上无法购买的单包装，其价格在不同城市间的分散度更高。而对于线上的多包装，由于天猫超市采取统一的定价策略，地区差异的解释几乎为0[[7]](#footnote-6)。

**表VIII1 价格方差分析（ANOVA）**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 商品集合 | 地区解释方差 | 时间解释方差 | 总方差 | 地区解释方差比例 | 时间解释方差比例 | 未解释的方差比例 |
| 线下 |  |  |  |  |  |  |
| 多包装 | 0.998 | 0.365 | 2.112 | 47.2% | 17.3% | 35.5% |
| 单包装 | 1.883 | 0.182 | 3.460 | 54.4% | 5.3% | 40.3% |
| 线上（天猫） |  |  |  |  |  |  |
| 多包装 | 0.053 | 2.767 | 3.087 | 1.7% | 89.6% | 8.7% |

为了得到价格分散度随时间变化的情况，我们对于线下多包装和单包装，分别计算了每个月的横截面IQR（Inter Quartile Range，四分位距），如图VIII2所示。多包装的地区间价格差异正在不断收窄，价格向统一趋近，然而这一趋势并没有出现在单包装中。



**图VIII2 线下价格分散度**

我们也尝试分析电商市场，是否降低了线下价格的地区间差异。我们对反事实模拟得到的价格序列，重复了上述相关统计。对反事实价格进行方差分析的结果如表VIII2所示。与原始统计量相比（见表VIII1），在两种反事实场景下，反事实价格被地区间差异解释的比例都有所升高，这意味着电商市场实际上能降低线下价格的地区间差异，使得线下价格趋向于收敛。

**表VIII2 反事实价格方差分析（ANOVA）**

|  |  |
| --- | --- |
| 反事实场景：天猫超市关闭 |  |
| 商品集合 | 地区解释方差 | 时间解释方差 | 总方差 | 地区解释方差比例 | 时间解释方差比例 | 未解释的方差比例 |
| 多包装（线下） | 0.922 | 0.173 | 1.665 | 55.4% | 10.4% | 34.2% |
| 单包装（线下） | 1.896 | 0.131 | 3.336 | 56.8% | 3.9% | 39.3% |
|  |  |  |  |  |  |  |
| 反事实场景：电商市场消失 |  |
| 商品集合 | 地区解释方差 | 时间解释方差 | 总方差 | 地区解释方差比例 | 时间解释方差比例 | 未解释的方差比例 |
| 多包装（线下） | 11.622 | 1.541 | 16.385 | 70.9% | 9.4% | 19.7% |
| 单包装（线下） | 5.599 | 0.293 | 8.188 | 68.4% | 3.6% | 28.0% |

图VIII3显示，如果天猫超市关闭，则相比于原始价格序列，多包装、单包装的价格分散度将有所上升，但可能由于存在其他电商收拢价格的效应，上升幅度不太明显；但如果电商市场消失，则多包装、单包装的价格分散度都将明显抬升，其中多包装的抬升效果尤为显著，甚至于高过了单包装。这意味着电商市场发展，实际上降低了地区间的价格分散度。



**图VIII3 反事实价格分散度**

参考文献

Chaudhuri, S., P. Goldberg, and P. Jia, “Estimating the Effects of Global Patent Protection in Pharmaceuticals: A Case Study of Quinolones in India”, *American Economic Review*, 2006, 96(5), 1477-1514.

Deaton, A., and J. Muellbauer, “An Almost Ideal Demand System”, *American Economic Review*, 1980, 70(3), 312-326.

**注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明附录下载出处**。

1. 2021年，申通快递市内续重约为1.5元/kg，顺丰速运、圆通快递市内续重约为2元/kg，韵达快递、中国邮政EMS市内续重约为3元/kg。（https://www.maigoo.com/goomai/179871.html） [↑](#footnote-ref-0)
2. 纯牛奶行业的代表性制造商蒙牛、伊利2019年销售毛利率分别为37.6%，37.4%；2019年90家零售上市企业的平均毛利率为31.9%，同年代表性电商平台京东、苏宁易购的毛利率水平，分别为14.6%和14.5%。企业毛利率数据均来自年报。 [↑](#footnote-ref-1)
3. 数据来自网经社电子商务研究中心的报告（http://www.100ec.cn/zt/2019wllsbg/）。 [↑](#footnote-ref-2)
4. 边际成本由(Ⅵ3)式的拟合值与边际成本冲击相加得到。天猫超市由于全国统一定价，因此倒推不出每个城市-月度的边际成本。但幸运的是，反事实的场景中天猫超市被关闭，我们并不需要知道其边际成本。 [↑](#footnote-ref-3)
5. 仅在不到1%的样本点上数值解不收敛，我们将这些数值解去除。 [↑](#footnote-ref-4)
6. 数学上IQR为第三四分位数与第一四分位数之差。显然IQR数值越大，变量分布越分散。 [↑](#footnote-ref-5)
7. 之所以不严格为0，主要是由于不同地区商品权重不同，以及同一商品订单达成时间不同（尽管都在同一月内）造成的。 [↑](#footnote-ref-6)