

全国第六届研究生数学建模竞赛



题目

我国城镇就业问题研究

摘要:

本文围绕我国城镇就业人数的影响问题,从当前经济社会发展现实出发,运用现实经济数据,建立了影响我国城镇就业人数的模型,并据此对 2009 年和 2010 上半年的就业情况进行预测。

对于问题一,我们选择城镇就业人数作为社会就业情况的衡量指标。再通过大量文献查阅,并结合经济学理论,首先确定了部分可能对城镇就业人数产生影响的因素,然后通过相关统计技术分析最终将居民消费价格指数,进出口总额, GDP 和科技投入确定为影响城镇就业人数的主要指标。

对于问题二,由于经济序列的非平稳性,我们对变量序列进行了单位根检验和协整检验,并据此建立了城镇就业人数和其影响因素的多元对数回归模型。考虑上述回归模型仅反映了变量间的长期稳定关系,我们借助 ECM 模型对变量间短期波动和长期稳定关系进行描述,在此基础上建立脉冲响应函数,并进行了方差分解。

对于问题三,由于在问题二建立的模型使用的是全国平均数据,为了避免这种平均化处理造成的数据丢失,从而建立更加精确的模型,我们将省作为地区划分的标准,建立分地区的城镇就业人数影响模型。为了不失一般性,我们又分别从东部,中部,西部各取一个样本构建面板数据,并据此建立了分地区的城镇就业人数影响模型。

对于问题四,在问题二建立的模型基础上,我们建立了向量自回归模型 (VAR),并结合当前经济运行数据,对 2009 年和 2010 年上半年的就业状况进行预测,得出 09 年大约新增就业 1040 万。

对于问题五,我们从建立的模型和就业前景仿真结果,提出了提高城镇就业人口数的三项政策建议:(1)继续保持国内生产总值的高速发展;(2)积极发展高质量对外贸易;(3)合理促进高新技术产业发展。

[关键词]: 城镇就业人数 协整 ECM 模型 脉冲响应 方差分解 VAR 模型

参赛队号 1035302

参赛队员 董挺 毕磊 许福娇

参赛密码 _____

(由组委会填写)

一、问题的重述

就业（或者失业）是社会、国民经济中极其重要的问题。从经济学的角度，影响就业（或者失业）的因素很多。从宏观层面上，消费、投资、政府购买和进出口都是重要的因素；而从中观层面，不同地区、不同产业也会表现出不同的特征。当然，中央政府调整宏观经济政策（包括财政政策和货币政策），以及对不同地区和不同产业实行不同的扶持政策都会对就业产生巨大的影响。

就我国的现实情况，2008年我国经济社会经受了历史罕见的考验，GDP依然保持9%以上平稳较快增长，城镇新增就业1113万人，城镇登记失业率为4.2%。2009年我国就业面临更大的挑战，一是国际金融危机导致国际市场需求难以在短期内复苏；二是今年我国经济增速下滑；三是国内消费需求乏力；四是一些行业产能过剩与市场预期不确定导致企业投资不足，所以就业形势十分严峻。

中央政府从08年10月开始实施了40000亿元的投资计划，确定了十大产业振兴计划，采取扩大国内消费需求的措施，提高对外开放水平以增加出口。同时，中央财政拟投入420亿元资金实施积极的就业政策。09年我国在就业方面的目标：城镇新增就业900万人以上，城镇登记失业率控制在4.6%以内。

请在参考就业问题的研究成果，利用近年来我国有关的统计数据并结合一年多来我国国民经济的运行数据，就我国就业人数或城镇登记失业率研究如下问题。

1. 对有关统计数据进行分析，寻找影响就业的主要因素或指标。
2. 建立城镇就业人数或城镇登记失业率与上述主要因素或指标之间联系的数学模型。
3. 对上述数学模型从包含主要的经济社会指标、分行业、分地区、分就业人群角度，尝试建立比较精确的数学模型。（由于时间限制，建议适度即可）
4. 利用所建立的关于城镇就业人数或城镇登记失业率的数学模型，根据国家的有关决策和规划对2009年及2010年上半年的我国就业前景进行仿真（可以根据模型的需要对未来的情况作适当的假设）。
5. 根据所建立的数学模型和仿真结果，对提高我国城镇就业人口数或减少城镇登记失业率提出你们的咨询建议。

二、问题的假设

- 1、假设不考虑劳动力供给对城镇就业人数的影响
- 2、假设统计数据真实可靠
- 3、假设 2009 年到 2010 年期间无突发事件影响

三、符号的说明

符号	符号说明
<i>JYRS</i>	全国城镇就业人数
<i>XFZS</i>	全国消费者物价指数
<i>JCK</i>	全国进出口总额
<i>GDP</i>	全国 GDP 总量
<i>KJTR</i>	全国科技投入总量
<i>SJYRS</i>	省城镇就业人数
<i>SAGDP</i>	省平均 GDP
<i>SJCK</i>	省进出口总额
<i>SXFZS</i>	省消费者物价指数

四、问题的分析及模型的建立及求解

4.1 问题的背景资料

充分就业一直是市场经济国家追求的三大宏观经济目标之一。学术界对此也进行了大量的理论和实证研究。目前，理论界基本上形成了以凯恩斯的非充分就业均衡理论，经济增长理论，发展经济学以及制度经济学就业理论等为代表的八大就业理论。而在实证方面，国内外学者也利用现实数据对失业问题进行了大量的实证研究。

近年来，我国经济保持了年均 8% 以上的高速增长，但就业问题依然严峻，现实中大量大学生就业难，社会中大量隐性失业等矛盾突出。特别是 08 年金融危机以来，外贸需求急剧萎缩，依靠外贸拉动经济的发展模式受到严重挑战，经济结构性矛盾日益突出，随之而来就是强大的就业压力。党和国家也采取了大量的措施以缓解当前的这一系列矛盾。并提出 09 年城镇新增就业 900 万人以上，城镇登记失业率控制在 4.6% 以内的就业目标。就业问题始终是经济社会发展长期不可回避的话题。也只有对我国就业问题进行了深入的分析，找出问题所在，才能提出好的对策破解我国就业难题。

本文正是在这一经济，社会背景下，对我国就业问题进行一定的探究，希望能对解决我国就业起一定帮助作用。

4.2 问题一的求解

4.2.1 问题一的分析

学术界对失业问题做了大量的研究，并提出了许多影响失业的因素。其中，奥肯定理认为失业率与 GDP 之间存在固定的比例关系，而“失业—物价”菲利普斯曲线认为失业率与通货膨胀率之间的存在着此消彼长的关系。另外还有许多其他理论对影响失业的因素进行了分析，提出了很多的指标。因此要研究就业问题，我们首先要根据相关理论和文献找出可能影响我国城镇失业问题的相关因素。然后，借助统计工具，对这些因素进行相关的统计分析，对这些因素进行筛选，并最终确定本文的研究指标。

4.2.2 问题一的求解

4.2.2.1 选择指标

通过对相关文献分析发现，由于我国经济具有特殊的二元结构，在农村劳动力市场的大量的隐形失业人群不断向城镇转移的过程中，城镇劳动力供给趋于无穷。对我国城镇劳动力市场而言，劳动力供给不是影响就业的主要因素，劳动力市场的需求方面对市场的均衡起着决定性作用。因此要研究就业问题也主要是研究劳动力市场的需求方面。

劳动力作为一种生产要素，其需求是由社会总需求所决定。所以对就业问题的研究也就更多的转化为对社会总需求的探究。而对社会总需求主要是从总量和结构两方面进行。总量方面，根据宏观经济理论，影响全社会需求总量的因素分别为投资，需求，出口。而对于结构方面，长期以来我国就存在经济结构发展不平衡的问题，而这也必然导致就业的不平衡，在传统产业中失业人口大量存在的同时，在某些领域则同时出现了劳动力紧缺的现象。同时随着经济结构不断调整，产业升级不断升级，经济结构的不平衡也不断加剧，这也将导致就业的结构性矛盾日益突出。具体而言，我们选择了以下指标作为研究的对象。

(1) 就业状况指标

评价社会就业状况的指标有多种，常见的有失业率和就业人数，考虑失业率是一个比例指标，同时受到城镇失业人数和城镇劳动力供给总量的影响。我国统计数据的局限性又导致城镇劳动力供给指标的缺失，这给问题的研究带来了不便，而城镇就业人数作为一个数量指标，本身只受到外部因素的影响，而且也能直观的反应社会就业状况。为了研究问题的方便，本文选择城镇就业人数作为城镇就业状况的指标。

(2) 影响城镇就业状况的指标

由于我国劳动力市场更多由劳动力的需求方面所决定，而劳动力的需求则由社会总需求所决定。所以我们从宏观经济理论出发，结合实际数据情况，确定了以下反应社会需求总量的指标作为影响城镇就业人数的因素：全社会固定资产投资总额，居民消费品零售总额，进出口总额，货币量 M1 和消费者物价指数。

另外，因为我国结构性失业问题大量存在，我们还需要从社会经济结构层面来考虑失业问题，对此我们确定了将三产/二产的比例，科技投入合计这两个对社会经济经济结构的指标作为影响我国城镇就业问题的因素。

最后，考虑到国企改革对我国经济的重要影响，我们加入了国企改革这一制度变量。最终我们初步确定了 10 个影响就业问题的因素。

对于反映制度变化的虚拟变量 D 定义如下：

$$D = \begin{cases} 1 & (1995\text{年国企改革后}) \\ 0 & (1995\text{年国企改革前}) \end{cases}$$

4.2.2.2 指标筛选

对于上述初步建立的指标体系，虽然在理论都可以找到支撑的依据，但其实际情况如何还有待检验，另外对于建立经济模型，过多的指标也限制了模型的实用性。下面我们就将利用简单的统计分析从上述 10 个影响因素中筛选出更加有效的影响因素。为了进行统计分析，我们从统计年鉴和相关财经网站上找出了上述 11 个指标从 1978 年到 2008 年以来 31 年的年度统计数据。并对此进行相关的统计检验。首先，我们对上述 11 个变量间的相关性进行检验，得到了各因素对城镇就业人数的相关系数如下表 2

表 1 各因素与就业人数的相关系数

	GDP	XFZS	SHLS	JCK	GDTZ	HB	PJGZ	KJTR	CYBL	<i>D</i>
JYRS	0.96	0.92	-0.04	0.96	0.96	0.97	0.91	0.95	0.94	0.78

从上表 1 中可见，上述 10 个因素除了社会零售总额外其他 9 个因素与城镇就业人数的相关系数较小都很高，大部分在 0.9 以上。这充分说明了各指标与就业人数存在较强的相关关系，也说明我们选择上述因素对就业人数的影响进行研究具有一定合理性。另外，社会零售总额与城镇就业人数相关性不高也不能因此就将其排除，因为相关系数反映的只是一种线性关系，没有线性相关，但可能存在非线性的关系。

在上述经济因素与城镇就业人数之间存在较大相关性同时，各指标之间相关系数也较大反映了各因素之间可能存在着线性相关。对于存在着线性相关的问题，由于其中的变量能够被其他变量所替代，如果直接建立模型可能会带来不好影响。对此，我们考虑将上述的因素进行进一步的缩减。首先，我们对上述各因素进行共线性诊断，并通过逐步回归和显著性检验对上述因素进行筛选。

表 2 各指标多重共线性诊断

系数	参数估计	估计标准差	VIF
β_1	0.40445	0.10873	2049.78977
β_2	0.08028	0.05154	10.99801
β_3	-0.33845	0.11255	1428.32900
β_4	0.08081	0.05120	623.12363
β_5	-0.18664	0.09727	3.87458
β_6	0.12405	0.03310	356.47369
β_7	-0.13997	0.06851	533.87058
β_8	-0.10366	0.05161	651.06365
β_9	-0.10910	0.08933	1086.47561
β_{10}	-0.08205	0.02633	15.65152

从表 2、表 3 输出结果可以看出，这些指标间具有明显的多重共线性。首先，最小的特征根为 0.00001534， $\phi_1 = 807.999$ ，它对某些系数方差的膨胀起到了决定性的影响。再看表 4，其中最小特征根对系数 β_3 和 β_7 的方差贡献率比为 80.4%，56.8% 因此估计的 β_3 和 β_7 将十分不稳定，结果很差，这一点从其对应的方差膨胀因子 VIF 为 1428.32900 和 533.87058 远超过了 10。综上所述，上述指标间存在较强的共线关系。

表 3 方差贡献比

特征根	ϕ_j	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
10.013	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.843	3.446	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013
0.132	8.722	0.000	0.053	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.155
0.010	32.275	0.000	0.389	0.001	0.000	0.000	0.003	0.000	0.001	0.000	0.164
0.001	86.396	0.004	0.108	0.001	0.000	0.002	0.036	0.027	0.000	0.015	0.007
0.000	172.200	0.011	0.054	0.005	0.249	0.004	0.220	0.001	0.066	0.002	0.204
0.000	184.998	0.013	0.079	0.000	0.014	0.002	0.605	0.002	0.259	0.003	0.025
0.000	215.093	0.058	0.008	0.002	0.117	0.004	0.048	0.025	0.127	0.155	0.126
0.000	508.505	0.655	0.001	0.073	0.138	0.202	0.045	0.160	0.038	0.295	0.036
0.000	542.509	0.001	0.273	0.113	0.481	0.774	0.019	0.218	0.116	0.224	0.258
0.000	807.999	0.258	0.022	0.804	0.000	0.011	0.024	0.568	0.393	0.306	0.012

对于上述因素存在的多重共线性问题，我们下面将采用逐步回归法，并运用显著性检验，对上述因素进行取舍，最后保留了居民消费价格指数，进出口总额，GDP 和科技投入作为影响城镇就业人数的最终指标。具体如表 4 所示。

表 4 影响城镇就业人数的因素

因素名称	符号
全国消费者物价指数	<i>XFZS</i>
全国进出口总额	<i>JCK</i>
全国GDP总值	<i>GDP</i>
全国科技投入总量	<i>KJTR</i>

4.3 问题二模型的建立与求解

4.3.1 问题二的分析

对于要建立上述 4 个经济因素指标与城镇就业人数的数学模型，我们首先想到进行回归拟合。但考虑到本文中的各经济社会指标本身都是一个时间序列，对于经济时间序列通常是非平稳的，根据计量经济学知识，对非平稳序列进行回归分析，可能会导致伪回归现象。所以，在建立回归模型前，我们首先要对序列的平稳性进行检验，如果序列平稳则可以直接进行回归拟合，而对于非平稳的序列则还需要根据实际情况进行协整分析，并在此基础上建立相应的回归模型。下面我们就按照这一流程，建立居民消费价格指数（XFZS），进出口总额（JCK）科技投入（KJTR），GDP 和城镇就业人数的数学模型

4.3.2 问题二模型的建立

4.3.2.1 单位根检验

为了避免因为上述经济变量的非平稳性产生的伪回归，我们首先要采用单位根检验法来对数据的平稳性进行检验。一般来说单位根检验有 ADF 法和 PP 法，本文采用 ADF 法对上述影响就业人数的各指标的时间序列进行单位根检验，其结果如表 5 所示。

表 5 单位根的 ADF 检验

变量	系数对应的 t 统计值	10%临界值	变量	系数对应的 t 统计值	10%临界值
GDP_t	-3.221728	-3.153936	ΔGDP_t	-3.233456	-3.248562
$KJTR_t$	-1.678836	-3.218382	$\Delta KJTR_t$	-2.622989	-4.003216
JCK_t	-1.678836	-3.218382	ΔJCK_t	-4.263979	-2.622989
$XFZS_t$	-0.096896	-1.609798	$\Delta XFZS_t$	-5.160149	0.0003
$JYRS_t$	-2.304618	-2.621007	$\Delta JYRS_t$	-4.913188	-2.622989

从上表 5 可以看出，所有指标的序列 JYRS, GDP, KJTR, JCK, XFZS 均为非平稳性序列，但其一阶差分序列在置信水平 10% 下都通过了显著性检验，这表明其均为平稳性序列。故上述五个序列都是一阶平稳的，也就是说这五个变量都是 I(1) 的过程。对于具有同阶单整序列，我们就可以对其进行协整分析。下面我就着手对这五个序列协整关系进行检验。

4.3.2.2 协整检验

上述单位根检验表明 JYRS, GDP, KJTR, JCK, XFZS 序列都是一阶单整序列，对于同阶单整序列其可能存在协整关系。协整模型的设定和滞后期确定原则是现根据最小化 AIC 和 SC 信息的标准选取，然后进行模型检验，若不能通过检验，则重新进行设定，直到找到相对理想的模型，经过反复检验确定协整变量具有线性趋势并有截距项。我们用 Johansen 最大似然法分析各个变量的协整关系，并求得协整向量，其结果整理如下：

表 6 各变量之间的协整检验结果

协整秩 H_0	最大特征根	迹统计量	5%的临界值	P值
$r \leq 0^*$	64.94610	135.4076	79.34145	0.0000
$r \leq 1^*$	31.35717	70.46148	55.24578	0.0013
$r \leq 2^*$	21.70105	39.10431	35.01090	0.0173
$r \leq 3$	15.71251	17.40327	18.39771	0.0685
$r \leq 4$	1.690758	1.690758	3.841466	0.1935

注：* 表示在 5% 显著性水平上拒绝原假设

表中 r 表示协整关系的个数，在 5% 的显著性水平拒绝 $r \leq 2$ 的假设，即变量之间存在三个协整关系。

4.3.2.3 对数回归模型

上述协整检验表明 JYRS, GDP, KJTR, JCK, XFZS 这 5 个变量之间存在一种长期的稳定关系，对此我们可以通过对其进行多元回归拟合，建立反映其相互联系的数学模型。我们首先尝试建立以 JYRS 为因变量，其他四个变量为自变量的多元线性回归方程，但模型通不过显著性检验。后来我们对数据进行了处理，并采取逐步回归法，但遗憾的是得到的模型拟合效果始终不理想。

对此，我们选用对数线性回归模型（非线性），试图对上述变量间的数学关系进行拟合。

对于多变量对数模型一般形式如下：

$$\ln(y) = \beta_0 + \sum_{i=1}^k \beta_i \ln(x_i) + u$$

$$\beta_i = \frac{\partial \ln(y)}{\partial \ln(x_i)} = \frac{\partial y / y}{\partial x_i / x_i}$$

其中， y 为因变量， x_i 表示第 i 个自变量， β_0 为截距项， β_i 为各自变量系数， u 为残差项。 β_i 称为偏弹性，它度量了在其他变量不变的条件下，解释变量 x_i 对因变量 y 的弹性影响。

结合本文所研究的问题，我们试图建立以 $\ln(JYRS_t)$ 为因变量，以 $\ln(GDP_t)$ ， $\ln(XFZS_t)$ ， $\ln(KJTR_t)$ ， $\ln(JCK_t)$ 为自变量的如下对数多元线性回归模型：

$$\ln(JYRS_t) = \beta_0 + \beta_1 \times \ln(GDP_t) + \beta_2 \times \ln(XFZS_t) + \beta_3 \times \ln(KJTR_t) + \beta_4 \times \ln(JCK_t) + u_t$$

其中 $\ln(JYRS_t)$ 表示第 t 年全国城镇就业人数， $\ln(GDP_t)$ 表示第 t 年全国 GDP， $\ln(XFZS_t)$ ，表示第 t 年全国消费价格指数， $\ln(KJTR_t)$ ，表示第 t 年全国科技投入总量， $\ln(JCK_t)$ 表示第 t 年全国进出口总额。

4.3.3 问题二模型的求解

我们选用 $\ln(JYRS_t)$ ， $\ln(GDP_t)$ ， $\ln(XFZS_t)$ ， $\ln(KJTR_t)$ ， $\ln(JCK_t)$ 五个指标从 1978 年到 2008 年 31 年的年度数据，运用 SAS 软件进行多元线性回归，其结果如下：

$$\ln(JYRS_t) = 6.67 + 0.44 \times \ln(GDP_t) - 0.5 \times \ln(XFZS_t) - 0.25 \times \ln(KJTR_t) + 0.22 \times \ln(JCK_t)$$

(11.53927) (1.942411) (6.656232) (8.726931) (-4.427182) (2)

$$R^2 = 0.986149, F = 462.7735$$

从 R^2 大小来看模型的拟合优度达到 98.6%，而且 F 统计值也通过了显著性检验，各系数也通过了 t 检验的显著性检验。因此建立的模型比较科学和合理。

对上述所建模型进行进一步分析发现：方程（2）中 $\ln(gdp_t)$ 和 $\ln(jck_t)$ 前面系数均为正数，说明 GDP 和进出口额对就业人数具有正的影响，且对应的偏弹性分别为 0.44 和 0.22。从经济学上意义是一单位 GDP 增长和一单位进出口额增长分别带来 0.44 单位和 0.22 单位就业人数的增加。而， $\ln(xfzs_t)$ 和 $\ln(kjtr_t)$ 前系数为负数，这表明消费者物价指数和科技投入对就业人数有负的影响。同样，一单位科技投入增长和一单位消费物价指数的增长分别带来 0.25 单位和 0.5 单位就业人数的减少。下面我们将对模型的经济意义进行进一步的剖析。

对于方程（2）中反映的 GDP 和进出口额对就业人数具有正的影响与实际也是相符的，GDP 和进出口的增长表现为经济的繁荣，必然带来劳动力需求的增加，所以能提高城镇就业人数。同时对比 GDP 和进出口对就业的偏弹性发现，进出口对就业人数的弹性达到 GDP 的一半，这也与我国 GDP 中外贸出口占有比例相符，反映了我国当前经济现状。另外，对于科技投入会导致城镇就业人数减少的结论也好理解。因为科技投入增加，带来劳动生产率的提高，而这会导致产业结构升级转型，传统的劳动密集型产业向资本和技术密集型转移，全社会对劳动力需求也将会减少，因此对科技投入的增加导致的科技进步和产业转型虽然有利于经济，社会长期发展，但同时也会产生失业增加等社会问题。

对于消费者物价指数与就业人数的反向变动则与传统经济学理论不符合。“失业—物价”菲利普斯曲线认为通货膨胀率和失业率是一种此消彼长的关系，即两者之间存在负相关关系。而本文所得结论—消费者物价指数（通货膨胀）与就业人数之间的这种负相关关系显然与传统理论矛盾。对于这种矛盾，我们对此做出的解释是，由于我国是一

个二元经济，面临大量农村劳动力向城镇转移导致的城镇劳动力供给趋于无穷的现状与菲利普斯曲线反映的经济基础不相同，导致了该曲线揭示的理论不适用于我国的现状。

4.3.4 对问题二模型的改进

4.3.4.1 误差修正 (ECM) 模型

经过前文的变量之间的协整分析可知，上述模型中的时间序列向量之间是协整的，也就是说从长期来看变量之间是具有均衡的关系。但上述建立的对数多元回归模型只提供变量间长期关系的信息，短期中由于会受到随即干扰的影响，这些变量有可能偏离均衡值，虽然这种偏离是暂时的，并且最终会回到均衡状态，但是这种短期动态的影响却无法从上述基于协整检验的对数回归模型中得到反映，正是基于上述模型的不足，我们对模型进行进一步的优化改进，对此，我们引入误差修正模型 (ECM) 模型，并借此研究各变量之间的长期稳定和短期动态变化的关系，然后通过脉冲响应来比较失业率对各变量一个单位正向冲击的反应结果，以及各变量对失业率的冲击效应分解，我们采用的统计软件是 Eviews6.0。

$$\Delta \ln(JYRS_t) = 0.157754 \times \Delta \ln(JCK_t) - 0.367954 \times \Delta \ln(XFZS_t) - 0.454400 \times ecm_{t-1}$$

$$(3.350575) \qquad \qquad (-3.681984) \qquad \qquad (-1.975246) \qquad \qquad (3)$$

其中误差修正项：

$$ecm_{t-1} = \ln(JYRS_{t-1}) - 6.67 - 0.44 \times \ln(GDP_{t-1}) - 0.22 \times \ln(JCK_{t-1}) + 0.25 \times \ln(KJTR_{t-1}) + 0.5 \times \ln(XFZS_{t-1})$$

调整后的 $R^2 = 0.995347$, $F = 213.9106$

误差修正项系数为 -0.454400 ，符合反向修正原则，表明短期的非均衡状态逐渐向长期的均衡状态趋近。

从式(3)来看，GDP 和科技投入在短期内并不会对就业人数产生显著的影响。当期进出口对就业人数有正向作用，相关系数为 0.157754 。当期的消费者指数对城镇就业人数存在负向作用，相关系数为 -0.367954 。由此可见，就业人数的短期变化可以分为两部分：一部分是短期进出口和消费者物价指数变动的影响；一部分是偏离长期均衡的影响。误差修正项的系数大小反映了对偏离长期均衡的调整力度。从系数来看，当短期波动偏离长期均衡的时候，将以 -0.454400 的调整力度将非均衡状态拉回到均衡状态。

4.3.4.2 脉冲响应函数

分析当一个误差项发生变化时，或模型受到某种冲击时对系统的动态影响的方法称之为脉冲响应函数法。脉冲响应函数就是用于衡量这种来自随机扰动项的一个标准差冲击对变量当前和未来取值的影响轨迹，它能够比较直观的刻画出变量之间的交互作用及其效应。下面我们就以 ECM 模型为基础，采用正交化方法和 Choleski 分解技术，建立就业人数与其他影响因素的脉冲相应函数模型。图 2 是基于 ECM(2)模型的城镇就业人数对四个影响因素的脉冲响应函数曲线。

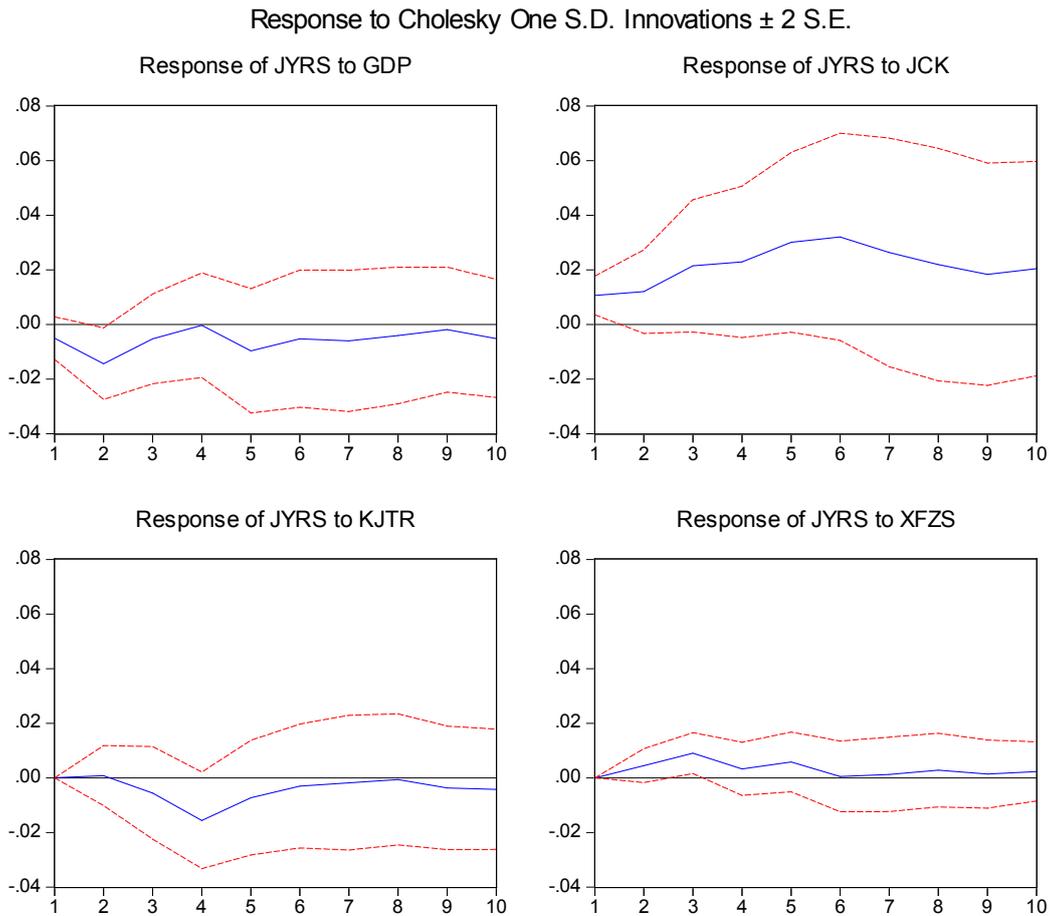


图 1 4 个变量各自的冲击对城镇就业人数的影响程度

图 1 分别给出了各变量一个标准差大小的冲击对城镇就业人数的脉冲响应函数，图中横轴表示冲击作用的期间数（年），纵轴表示城镇就业人数的变换程度，曲线表示了脉冲响应函数，代表了各相应变量冲击的动态响应；两侧的虚线是脉冲响应函数加减两倍标准差的值，表明冲击响应的可能范围。

图 1 是 4 个影响城镇就业人数变量的冲击引起的城镇就业人数变化的脉冲响应图。我们逐一对其进行分析。

从图中可以看出，当在本期给 GDP 一个标准差的冲击后（即增加），影响就业人数先减少后上升，并在第 4 期到达最高，此后趋于稳定，并在第 6 年以后收敛于 0，这说明 GDP 对就业人数的影响的响应大约持续为 6 年左右。

进出口的冲击对城镇就业人口的影响则是一个长期的过程，从图中可以发现，当本期给进出口一个标准差的冲击后，其影响城镇就业人数持续上升，即使在 10 年后还保持较高水平。

科技投入对城镇就业人数的冲击最明显，其对城镇就业人数的减少在第 4 年处于最大水平。此后也逐步收敛于 0。这说明科技投入后在第四年左右对城镇就业人数的负面影响最大程度爆发。

而居民消费指数对城镇就业人数的冲击不明显。

4.3.4.3 方差分解

我们可以使用方差分解法通过求解扰动项对向量自回归模型预测均方误差的贡献

度，了解各类因素对就业人数的冲击作用，各个变量的方差分解结果见表。

表 7 *JYRS* 的方差分解

时期	标准差	GDP	JCK	JYRS	KJTR	XFZS
1	0.028558	5.913390	25.58006	68.50655	0.000000	0.000000
2	0.045865	26.78269	29.36215	41.50608	0.072324	2.276752
3	0.062458	17.19454	46.99485	27.07077	2.060717	6.679121
4	0.083225	11.33552	53.66962	18.20401	11.92574	4.865107
5	0.112749	10.48070	63.09574	12.45665	9.665749	4.301157
6	0.148502	8.416641	69.55771	11.40658	7.407965	3.211108
7	0.183791	7.804283	71.73590	11.38226	6.332348	2.745204
8	0.221115	7.321681	72.76076	11.58675	5.720751	2.610051
9	0.256646	6.920015	73.47173	11.55586	5.577558	2.474829
10	0.283504	6.814936	74.26993	11.09665	5.442771	2.375705

从表 7 中可以看出，随着中国的对外贸易量的扩大，进出口总额对就业人数的解释力度不断加大。长期而言除了就业人数自身的变动影响外，GDP 和科技投入值是影响就业人数最重要的因素，分别解释 6.8%和 5.4%的就业人数的变动。

4.4 问题三模型的建立与求解

4.4.1 问题三的分析

由于问题二中建立各因素与城镇就业人数的数学模型中所用的数据是全国数据。全国数据作为一个平均值，没有充分反映个体的差异性。为了建立更加精确的各因素与就业人数的数学模型，我们就应该避免全国数据对数据平均化处理后造成数据信息丢失问题。为此，我们需要针对单个地区，或行业就业数据进行分析，建立针对地区或行业的城镇就业人数与相关影响因素的数学模型。

4.4.2 问题三模型的建立

为了建立更加精确的城镇就业人数与相关因素的精确数学模型，考虑到时间的问题，我们仅分地区建立了城镇就业人数的数学模型。

考虑到数据的可得性和可操作性，我们选择以省作为划分地区的标准。同时，为了不失一般性，我们按照全国东部，中部，西部的划分方式，分别在东部，中部，西部各选择一个省份，作为研究的对象。最后我们选择的省份或直辖市分别为：上海市，湖北省，四川省。同时考虑到统计口径的差异，我们将上述指标作了相应调整。新构建指标体系包括：省城镇就业人数 (*SJYRS*)；省平均 GDP (*SAGDP*)；省进出口总额 (*SJCK*)；省消费者物价指数 (*SXFZS*)。结合本文讨论问题，试图建立如下对数线性回归模型。

$$\ln(SJYRS_{it}) = \beta_0 + \beta_{1i} \times \ln(SAGDP_{it}) + \beta_{2i} \times \ln(SJCK_{it}) + u_{it}$$

其中 *SJYRS_{it}* 表示 *i* 省第 *t* 年就业人数。*SAGDP_{it}* 表示 *i* 省第 *t* 年人均 GDP 水平，*SJCK_{it}* 表示 *i* 省第 *t* 年进出口总额。*u_{it}* 为残差项，*i* = 1,2,3 分别对应上海市，湖北省和四川省。

数据上，我们选取了上海市，湖北省和四川省从 1995 年到 2008 年 14 年，上述 3

个变量的年度数据进行拟合。

4.4.3 问题三模型的求解

首先笔者对数据进行对数化处理,将处理后的数据,以估计截面残差的方差为权数,然后采用广义最小二乘法 (GLS),对此面板数据进行分析,通过 EVIEWS 软件运行相关程序,最后得到结果如下:

上海市:

$$\ln(sjyrs_{it}) = 3.273361 + 0.3173168 \times \ln(sagdp_{it}) + 0.078292 \times \ln(sjck_{it}) + 1.532427 \times \ln(sxfzs_{it}) + \alpha_i^*$$

t 值 (493.15) (142.043) (23.52090) (52.46667)

湖北省:

$$\ln(sjyrs_{it}) = 3.273361 + 0.371891 \times \ln(sagdp_{it}) + 0.294875 \times \ln(sjck_{it}) + 0.772740 \times \ln(sxfzs_{it}) + \alpha_i^*$$

t 值 (493.1503) (69.48919) (108.25380) (99.42362)

四川省:

$$\ln(sjyrs_{it}) = 3.273361 + 0.041655 \times \ln(sagdp_{it}) + 0.091387 \times \ln(sjck_{it}) + 2.537304 \times \ln(sxfzs_{it}) + \alpha_i^*$$

t 值 (493.1503) (139.1834) (33.15992) (9.744341)

表 8 各地区的不同截距

上海 α_i^*	湖北 α_i^*	四川 α_i^*
-1.382000	1.007526	0.374474

4.4.3 问题三模型的检验

首先是经济学上的检验,也就是定性的观察各个指标是不是与实际情况相符合的,我们发现人均 GDP,进出口总额,以及居民消费价格指数对就业人数是正指标,所以它们的前面系数是正值,符合实际。

其次,我们对其参数进行检验,有上述公式里的 T 统计量都显著的大于 2,所以认为这几个指标通过 t 检验,也就是说参数是显著的。对于模型的检验,我们则要看 F 统计量,从下表我们看出 F 值为 524.8577,也就是说模型也是合理的。也由此有了我们以下的分析。

表 9 统计检验结果

加权后的统计值			
R^2	0.994831	均值	3.915375
调整 R^2	0.992935	标准差	1.445345
F 统计值	524.8577	AIC	3.409866
残差平方和	0.023521	SC	3.906343
极大似然估计值	-59.60718	HQC	3.591844
没有加权的统计值			
R^2	0.995088	均值	3.330301
残差平方和	0.023521	DW	1.325115

(1) 从上表中可以发现,每一个地区截距项以及指标系数都不同,代表每个地区对就业人数的影响不同。由于各项指标是经过对数化处理的,所以其系数就是就业弹性。所谓就业弹性,就是经济增长(其他指标)每变化一个百分点所对应的就业数量变化的百分比。就业弹性为正值时:弹性越高经济增长(其他指标)对就业的拉动效应越大,反之则拉动效应越低。就业弹性为负值时,经济增长(其他指标)对就业的作用可以分为

两种:一种为“挤出”效应,是指经济为正增长但就业减少的情况,此时就业弹性绝对值越大对就业“挤出”越大;另一种为“吸入”效应,这种效应是指经济(其他指标)为负增长但就业增加的情况,此时就业弹性绝对值越大对就业的“吸入”效应就越大。

(2) 人均GDP对就业人数的影响是正向的,因为人均GDP提高,表明我国经济较景气,能够提供更多的岗位给劳动者,并且我国是人口大国,能够提供大量的劳动力,也就是说有岗位需求,因此,就业情况较好。但是上海,湖北等相对较发达的东、中地区前面的系数小于四川等西部地区,究其原因,可能是因为较发达的地区劳动生产率比较高,同样多的国内生产总值所需要的职工人数就相对的少,也就是能提供的岗位相对少,是比较符合实际的。

(3) 我们观察到进出口对就业人数的影响效应为负,进出口提高,表明对商品的需求增加,也就是说职工工作的岗位供给就相应的提高,从而吸纳剩余劳动力。但是它们的弹性有大有小(上海<四川<湖北),可能是因为进出口的商品一般指需加工的商品,吸纳的廉价劳动力。但是上海的资本密集型产业较多,而劳动密集型企业较少,并且工资价格水平也较高,所以即使有需要也往往运送中部加工,正好湖北等中部地区有大量廉价劳动力,并且离东部也较近,所以由于进出口吸纳的劳动力就越多,西部地区虽有大量劳动力,但是距离远,所以弹性又相对的少了。

(4) 最后居民消费价格指数与就业也是正向关系,也是符合实际的。我们可以结合第二类费利普斯曲线来说明这个现象。第二类费利普斯曲线描绘的是通货膨胀与失业之间的关系:即当失业率较高时,通货膨胀率便较低;而当失业率较低时,通货膨胀率便较高,也称为“失业—物价”菲利普斯曲线。那么在这里居民消费者价格指数对就业也有类似的表现,只不过两者的表示是正相关。

4.5 问题四模型的建立与求解

4.5.1 问题四的分析

在问题二中,我们建立了城镇就业人数与相应指标之间的对数回归模型,上述模型虽然能够对变量间相互关系进行很好的拟合,但其在对未来进行预测时则发生了困难。对于本问中要求根据国家的有关决策和规划对2009年及2010年上半年的我国就业前景进行仿真。很显然其实通过当前期数据对未来期数据进行预测,这样我们很自然想到了向量自回归(VAR)模型。

4.5.2 问题四模型的建立

由于VAR是一种非理论性的模型,它无需对变量做任何先验性约束,该模型采用多方程联立的形式,它不以经济理论为基础。在模型的每一个方程中,内生变量对模型的全部内生变量的滞后项进行回归,从而估计全部内生变量的动态关系。

通过向量自回归(VAR)模型法进行预测时,主要是依靠隶属数据的变化来预测未来数据,同时考虑相关变量的交互变化,其结合了回归分析和时间序列分析两种方法。通常,模型没有任何外生变量,每一内生变量都由它的滞后或者过去值以及模型中所有其他内生变量的滞后或过去值解释。对于一个包含 N 个变量的 P 阶向量自回归模型,记作 $VAR(P)$,表示为:

$$Y_t = c + BY_{t-1} + BY_{t-2} + \dots + BY_{t-k} + u_t$$

其中, $Y_t = (y_{1t}, y_{2t}, \dots, y_{Nt})'$, $C = (C_1, C_2, \dots, C_N)'$, u_t 为残差项。

根据本问研究实际,我们建立了如下关于城镇就业人数与其相关因素的的向量自回归(VAR)模型,具体形式如下:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

其中: $Y_t = (\text{JCK}_t, \text{JYRS}_t, \text{KJTR}_t, \text{XFZS}_t, \text{GDP}_t)'$, $c = (c_1, c_2, c_3, c_4, c_5)'$

$$\phi_i = \begin{pmatrix} \phi_{11}^i & \phi_{12}^i & \phi_{13}^i & \phi_{14}^i & \phi_{15}^i \\ \phi_{21}^i & \phi_{22}^i & \phi_{23}^i & \phi_{24}^i & \phi_{25}^i \\ \phi_{31}^i & \phi_{32}^i & \phi_{33}^i & \phi_{34}^i & \phi_{35}^i \\ \phi_{41}^i & \phi_{42}^i & \phi_{43}^i & \phi_{44}^i & \phi_{45}^i \\ \phi_{51}^i & \phi_{52}^i & \phi_{53}^i & \phi_{54}^i & \phi_{55}^i \end{pmatrix}, i = 1, 2, \dots, p$$

4.5.3 问题四模型的求解

利用问题二中相关经济统计数据, 我们建立 VAR 模型, 相关结果如下:

表 10 VAR(P)模型 AIC 和 SC 值结果

Lag	LR	AIC	SC	HQ
0	NA	-5.586875	-5.348981	-5.514148
1	286.2455	-16.81232	-15.38496*	-16.37596
2	49.33236*	-17.92851	-15.31168	-17.12852
3	28.14301	-18.48805*	-14.68175	-17.32442*

从上表可以看出, 根据 AIC 信息准则和 SC 信息准则值的最小化标准, 我们确定滞后阶数 P=2。对应于参数 c, 和变量 Yt, Yt-1 的系数矩阵估计值如下表 11, 表 12, 表 13, 其中括号内数据为估计参数的标准差。

表 11 变量 C 的矩阵

-27.45387	-5.690222	0.884226	5.755755	-4.767038
(9.26140)	(1.92901)	(6.28172)	(3.12562)	(2.85583)

表 12 Yt 系数矩阵

0.088658	-0.042624	-0.123001	-0.061463	-0.181170
(0.31698)	(0.06602)	(0.21500)	(0.10698)	(0.09774)
3.082179	0.877511	0.186407	0.383390	0.562141
(1.04436)	(0.21752)	(0.70836)	(0.35246)	(0.32204)
1.081985	0.099766	0.854999	-0.044122	0.173690
(0.36202)	(0.07540)	(0.24555)	(0.12218)	(0.11163)
1.374868	0.604854	-0.805940	0.257378	-0.272140
(0.88869)	(0.18510)	(0.60277)	(0.29992)	(0.27404)
0.099783	-0.630303	0.585387	0.679745	1.673011
(0.88300)	(0.18391)	(0.59891)	(0.29800)	(0.27228)

表 13 Y_{t-1} 系数矩阵

-0.015784	-0.020348	0.066350	0.152117	0.109801
(0.21850)	(0.04551)	(0.14820)	(0.07374)	(0.06738)
1.110966	0.548805	0.373357	-0.388069	0.596639
(1.32088)	(0.27512)	(0.89591)	(0.44578)	(0.40730)
0.140710	0.050925	0.134422	-0.000262	0.117683
(0.47290)	(0.09850)	(0.32075)	(0.15960)	(0.14582)
0.576895	0.224389	0.195282	-0.471160	0.243143
(0.79565)	(0.16572)	(0.53967)	(0.26852)	(0.24535)
-0.457059	0.526347	-0.570688	-0.777017	-0.988419
(0.77062)	(0.16051)	(0.52268)	(0.26007)	(0.23763)

根据上述建立的 VAR (2) 模型，我们对 1978 年以来城镇就业人数的实际值与模型预测值进行对照，其结果见下图 2

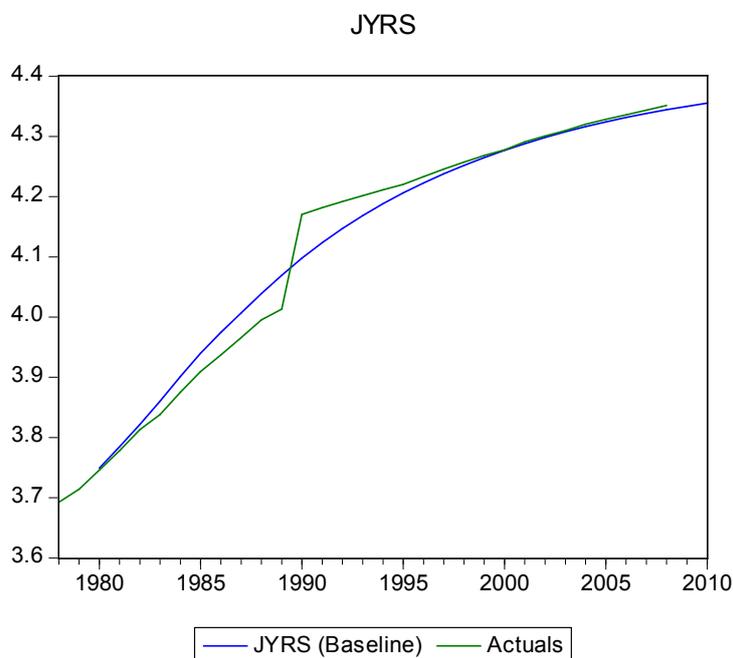


图 2 VAR 模型拟合效果图

从上图可以看出，上述建立的 VAR(2)能够很好的对历史数据进行拟合，我们利用上述模型对 2009 年和 2010 年的全国城镇就业人数进行预测，得到预测值分别为：2009 年的城镇就业人数为 31430 万人，2010 年的城镇就业人数为 32490 万人，也就是说 09 年新增城镇就业人数 1040 万，2010 年新增就业人数 1060 万。

虽然预测结果比政府报告设定目标高了 140 万个岗位，但我们切不可掉以轻心，在当前经济条件下，要实现年增 900 万就业岗位目标的压力也还是比较大的。特别是在目前全球经济还未完全走出金融危机阴影的背景下，虽然政府通过各项政策暂时避免了经济的进一步下滑，但同时也应该看到经济复苏的迹象还不明显，复苏的基础还不牢固。这就需要我们以本文建立的数学模型为依据，从对增加城镇就业人数有正向作用的指标出发，有针对性的开展工作，有效提高劳动力需求，努力实现政府报告提出的 09 年新增 900 万岗位的目标，并争取达到我们预测的结果。

4.6 问题五的分析

通过对本文建立的数学模型进行分析，我们发现从长期看 GDP，进出口总额对城镇就业人数有正向作用，而消费者物价指数和科技投入对城镇就业人数有负向影响。所以从长期来看，我们应该针对 GDP 和进出口总额，制定相应的措施，从而才能从根本上化解我国就业问题。具体而言应该从以下方面进行努力。

1. 继续保持国内生产总值的高速发展。

对我们这样一个农业人口占大多数的国家而言，只有保持高速的经济发展水平才能在发展中解决一系列经济、社会问题。更有专家提出，中国经济只有保持 7% 以上的增速才能不产生重大社会问题。由此可见，我国只有保持较快的经济发展，才能有效增加就业人数，减少城镇的登记失业率。为达到高的经济增长率，国家应实行积极的宏观经济政策，加大政府投资，鼓励消费扩大内需。

2. 积极发展高质量对外贸易。

从文中建立的模型，我们可以看出，在我国，外贸对就业人数的增加具有重要作用。对此我们需要从两方面理解。一方面，外贸如此重要，我们需要大力发展对外贸易，从而可以增加城镇就业人数。但同时我们还应该看到，近年来，我国对外贸易不断增长，外贸已成为拉动经济发展的重要力量。但这种发展模式有其局限性，经济过度的依赖外部经济，必然导致经济的非独立性。比如在当期，随着美国次贷危机导致的全球金融危机不断蔓延，西方国家传统的信用消费模式受到挑战，一向透支消费的“美国人”也开始储蓄的，与之相对的是我国外贸环境越来越严峻。在当前世界经济还没有完全走出金融危机阴影的前提下，贸易保护势力又有所抬头，美国对中国的轮胎特保案就是典型的例子。可以想象，在未来各国间贸易摩擦将会越来越多。正是因为我国过度依赖利用对外贸易来拉动经济发展，所以一旦金融危机爆发，世界经济陷入衰退，我国的失业人数将显著上升。所以，综合来看，我们不可一味的追求对外贸易数量的增加，更应该从质量上下功夫，努力提高外贸质量。

3. 合理促进高新技术产业发展

高新技术的发展是一把双刃剑，一方面，高新技术发展可以提高我国生产力水平，有助于增强我国的经济综合竞争力；另一方面，高新技术必然导致技术和资本对劳动的替代作用。随着传统的劳动密集型产业向资本密集型和技术密集型产业的不断转移，必然会带来失业增加等一系列社会问题。

五、模型的评价及推广

本文紧密围绕我国城镇就业人数的影响问题，从当前经济社会发展现实出发，运用现实经济数据，建立了影响我国城镇就业人数的模型，并据此对 2009 年和 2010 上半年的就业情况进行预测。

在问题一中，我们在对影响城镇就业人数的因素选择过程中，采用了文献查阅——指标初选——统计分析——指标筛选的方法。对于人文社科类的研究，因为往往最后结果是不确定，这就需要对前人基础上进行总结，并根据相关技术进行分析判别，所以本文在问题一的研究中所采用这种理论到实证的方法对于人文社科的其他研究也是适用的。

在本文问题二的研究中，我们先后对变量序列进行单位根检验，协整检验，并据此建立了城镇就业人数和其影响因素的多元对数回归模型。同时，在分析了回归模型局限性后，借助 ECM 模型对原模型进行了修正。对变量间短期波动和长期稳定关系进行描述，并在此基础上建立脉冲响应函数和方差分解。问题二中模型不仅适用于就业模型的建立，对于一般的经济问题，也可以按照这一流程进行分析，因此该模型具有良好的推广性。

问题三的模型主要是基于有效保留信息这一目标，通过分地区进行城镇就业人数与其相关因素的回归分析，通过选择多样本时间序列构造面板数据，并以此对变量间的相互关系进行拟合。面板数据作为截面数据和时间序列的结合，能够有效克服两者的不足，其运用也越来越广泛。

在问题四，我们建立了向量自回归（VAR）模型，由于 VAR 是一种非理论性的模型，它无需对变量做任何先验性约束，且该模型采用多方程联立的形式，在模型的每一个方程中，内生变量对模型的全部内生变量的滞后项进行回归。它的这些特点使得它在预测上具有良好的性能。

而由于时间仓促，本文在模型建立上还存在问题，另外，对于问题三，也因为时间原因，我们仅仅分地区进行了研究，而对分行业和分人群没有进一步分析，这也为我们未来的研究提供了一个发展方向。

[参考文献]:

- [1] 张平 等, MATLAB 基础与应用, 北京: 北京航空航天大学出版社, 2001 年
- [2] 苏金明, 张莲花 等, MATLAB 工具箱应用, 北京: 电子工业出版社, 2004 年
- [3] 曲庆云, 赵晓梅, 阮桂海等, 统计分析方法——SAS 实例精选, 北京: 清华大学出版社, 2005 年
- [4] 曾湘泉等, 面向市场的中国就业与失业测量经济学, 北京: 中国人民大学出版社, 2007 年
- [5] 袁乐平, 周浩明等, 失业经济学, 北京: 经济科学出版社, 2004 年
- [6] 韩兆洲, 安宁宁, 最低工资、劳动力供给与失业——基于 VAR 模型的实证分析, 广东: 暨南大学学报, 2007 年第 1 期
- [7] 黄新飞, 舒元, 基于 VAR 模型的 FDI 与中国通货膨胀的经验分析, 广东: 世界经济, 2007 年第 10 期
- [8] 高铁梅, 计量经济学分析方法与建模, 北京: 清华大学出版社, 2006 年
- [9] 赵博, 雍家胜, 菲利普斯曲线研究在中国的实证分析, 北京: 管理世界, 2004 年第 9 期
- [10] 张卫东, 汪海, 我国环境政策对经济增长与环境污染关系的影响研究, 湖北: 华中科技大学出版社, 2007 年
- [11] 葛洪伟, 王士同, 可能性线性模型中的参数选优与输入噪声间关系的研究, 江苏: 识别与人工智能, 2007 年 2 月
- [12] 袁志刚, 中国的乡—城劳动力流动与城镇失业: 一个经验研究, 北京: 中国分配就业论坛, 2006 年
- [13] 顾剑华, 政府公共支出对 GDP 长期增长效应的动态分析——基于广西数据的协整检验和 VAR 模型分析, 广西: 工业技术经济, 2007 年 11 月