

Qualidade Estatística do Inquérito ao Emprego

DEPARTAMENTO DE METODOLOGIA ESTATÍSTICA

Lucília Carvalho

SUMÁRIO

- **Componentes da Qualidade Estatística segundo o EUROSTAT**
- **Avaliação da qualidade estatística do Inquérito ao Emprego**
 - * **Precisão Estatística**
 - * **Coerência e Comparabilidade**
 - **A série longa do Inquérito ao Emprego**
 - **A nova base de Amostragem**
 - **Transição entre bases**



EUROSTAT

Working Group - **Assessment of Quality in Statistics**
Task Force - **Quality Indicators**

OCDE

One of the main objectives of the Statistics Strategy
OECD Quality Framework

FMI

Data Quality Assessment Framework



EUROSTAT

COMPONENTES DA QUALIDADE:

- 1 - Pertinência**
- 2 - Precisão**
- 3 - Actualidade e Pontualidade**
- 4 - Acessibilidade e Clareza**
- 5 - Comparabilidade**
- 6 - Coerência**
- 7 - Abrangência**

Carga estatística e Custo



COMPONENTES DA QUALIDADE DO EUROSTAT

1 - Pertinência

Refere-se à adequabilidade das estatísticas produzidas às necessidades dos utilizadores.

Envolve a a descrição dos tipos de utilizadores, das suas necessidades de informação e do nível a que estas se encontram satisfeitas.

A pertinência é difícil de avaliar sem recorrer a inquéritos aos utilizadores. Normalmente descrevem-se apenas os tipos de utilizadores e usam-se indicadores ad hoc, dos quais podem fazer parte, por exemplo, avaliações feitas por assembleias onde estão representados os utilizadores.

**COMPONENTES DA QUALIDADE DO EUROSTAT****1 - Pertinência no IE**

Tipos de utilizadores	Descrição
Internos	Departamento de Contas Nacionais Serviços de Estudos
Nacionais exteriores ao INE	Presidência do Conselho de Ministros –DPP Ministério das Finanças –GEE Ministério da Segurança Social e do Trabalho Ministério da Educação Banco de Portugal Instituto de Emprego e Formação Profissional Conselho Superior de Estatística Confederação Geral dos Trabalhadores Portugueses União Geral dos Trabalhadores Investigadores Universidades
Internacionais	Organização Internacional do Trabalho –OIT OCDE Banco Central Europeu



COMPONENTES DA QUALIDADE DO EUROSTAT

3 - Actualidade e Pontualidade

A Actualidade refere-se ao período de tempo que medeia entre a recolha de informação e a disponibilização dos resultados.

A Pontualidade refere-se ao intervalo de tempo que decorre desde a data prevista para a publicação dos dados e a sua disponibilização efectiva.

Envolve a fixação de indicadores com base na referência seguinte:

A informação referente a determinado período n deve ser disponibilizada antes do fim do período $n+1$.

Padrão de Qualidade - Período estabelecido por acordo com os responsáveis do projecto

Prazo de Disponibilidade - Período de tempo entre o fim do período de referência e a data de disponibilização dos resultados.

Desvio - Diferença entre o Padrão de Qualidade e o Prazo de Disponibilidade

**COMPONENTES DA QUALIDADE DO EUROSTAT****3 - Actualidade e Pontualidade do IE**

Período	Padrão de Qualidade n+ 50 dias	Desvio em dias
4º Trimestre de 2001	20 Fevereiro	-1
1º Trimestre de 2002	20 de Maio	+1
2º Trimestre de 2002	20 de Agosto	-10
3º Trimestre de 2002	20 de Novembro	+1



COMPONENTES DA QUALIDADE DO EUROSTAT

4 - Acessibilidade e Clareza

A Acessibilidade diz respeito às condições físicas segundo as quais os utilizadores podem obter os dados.

(onde se dirigir, como pedir, tempo de disponibilização, ...)

A Clareza refere-se ao contexto informacional dos dados.

(Se têm metainformação associada, se têm associada informação gráfica, se existe informação sobre a qualidade dos dados, limitações de utilização, ...)

O INE está a fazer um enorme esforço no sentido de melhorar esta componente da qualidade. Além de ter construído um protótipo de Documento Metodológico com a descrição pormenorizada de toda a meta-informação associada às Operações Estatísticas, está também a construir um protótipo de Relatório de Qualidade que deve acompanhar obrigatoriamente as ocorrências das Operações Estatísticas.



COMPONENTES DA QUALIDADE DO EUROSTAT

4 - Acessibilidade e Clareza do IE

Acesso aos dados é feito através:

- Destaque para a Comunicação Social e Publicação Trimestral;
- INTERNET (via INFOLINE) Pub. Trimestral e séries cronológicas;
- Quadros do plano de apuramentos trimestral
- Bases de microdados anonimizadas

Informação disponibilizada com os dados:

- Informação referente aos erros de amostragem
- Informação referente à metodologia, conceitos, ...



COMPONENTES DA QUALIDADE DO EUROSTAT

6 - Coerência

A Coerência refere-se à adequabilidade das estatísticas a uma segura comparação ou combinação de diferentes formas e com várias finalidades.

Envolve a descrição das razões de falta de comparabilidade, quando existe, entre a mesma estatística produzida por operações estatísticas diferentes; ou entre estatísticas diferentes incidindo sobre o mesmo período de referência.

A coerência é difícil de medir e o objectivo desta componente é, sobretudo, o de informar os utilizadores da informação estatística, das diferenças conceptuais existentes entre estatísticas, no que diz respeito a conceitos e definições, universo de referência, classificações e âmbito geográfico.



COMPONENTES DA QUALIDADE DO EUROSTAT

6 - Coerência do IE

Foi feita uma análise da coerência existente entre diversas variáveis apuradas no IE e noutras operações estatísticas:

- **Entre o IE e as Contas Nacionais**

Emprego

- **Entre o IE e o Inquérito às Empresas Harmonizado**

Emprego

Emprego por CAE

Nº de Horas Trabalhadas

- **Entre o IE e o Emprego Registrado**

Emprego



COMPONENTES DA QUALIDADE DO EUROSTAT

7 - Abrangência

A Abrangência diz respeito ao grau em que as estatísticas estão disponíveis, tendo como referência o Sistema Estatístico Europeu.

7 - Abrangência do IE

Todas as variáveis estão disponíveis.



COMPONENTES DA QUALIDADE DO EUROSTAT

2 - Precisão

A precisão refere-se à proximidade entre o valor obtido para uma característica da população e o seu valor real desconhecido.

Envolve uma avaliação dos seguintes erros:

- Erros amostrais
- Erros não amostrais
 - Erros de cobertura (base de amostragem)
 - Erros de Medida
 - Erros de Processamento
 - Erros de não-resposta



COMPONENTES DA QUALIDADE DO EUROSTAT

2 - Precisão do IE

ERROS AMOSTRAIS

- O Regulamento CE 577/98 fixa a dimensão do erro relativo da estimativa do desemprego médio anual, por NUTS II, e o erro relativo da variação do desemprego trimestral, a nível nacional.
- Os erros relativos (coeficientes de variação) são difíceis de calcular devido à complexidade do desenho amostral e envolvem técnicas relativamente recentes de cálculo de variâncias através de métodos de simulação como, por exemplo, o método de Jackknife.
- Os coeficientes de variação são calculados para todas as variáveis dos quadros de apuramento e figuram na publicação do IE.

**COMPONENTES DA QUALIDADE DO EUROSTAT****2 - Precisão do IE****ERROS AMOSTRAIS**

Taxa de Desemprego e Coeficientes de Variação									
	1t02		2t02		3t02		4t02		
Norte	3,8%	7,07	4,1%	6,34	5,4%	5,64	6,2%	5,61	
Centro	3,0%	13,71	2,8%	16,52	2,5%	16,13	3,7%	13,61	
Lisboa	5,9%	6,72	6,2%	7,83	6,4%	6,55	7,6%	6,86	
Alentejo	6,2%	12,32	5,1%	14,62	7,4%	13,12	8,0%	10,59	
Algarve	5,1%	13,43	4,5%	13,18	4,7%	12,53	6,9%	10,91	
Açores	2,7%	17,96	2,2%	20,88	2,7%	21,58	2,3%	18,50	
Madeira	2,6%	18,42	2,3%	19,66	2,7%	16,46	2,3%	18,30	
País	4,5%	4,18	4,5%	4,63	5,1%	3,87	6,2%	3,88	



COMPONENTES DA QUALIDADE DO EUROSTAT

2 - Precisão do IE

ERROS NÃO AMOSTRAIS

Erros de Cobertura - Erros que dizem respeito à base de amostragem. São provocados normalmente pela desactualização da base de amostragem e avaliam-se através do número de unidades da amostra que se verificou estarem fora do âmbito do inquérito ou em duplicado (Sobrecobertura) e também do número de unidades que num dado período se verificou não estarem listadas na base (Subcobertura).

Erro de cobertura de 2002 - 16% num total de 20 747 U.A.

O INE procedeu a uma actualização da base de amostragem (Amostra-mãe) nas subsecções da nova amostra do IE que se encontravam com uma elevada percentagem de moradas de difícil localização.



COMPONENTES DA QUALIDADE DO EUROSTAT

2 - Precisão do IE

ERROS NÃO AMOSTRAIS

Erros de Medida - Referem-se ao registo errado de informação e são medidos através de taxas de consistência de respostas e taxa de erros de codificação.

O INE procede, todos os trimestres, a um **Inquérito de Qualidade** incidindo sobre cerca de 10% das unidades de alojamento da amostra.

Este inquérito é efectuado por um supervisor que reentrevista os alojamentos escolhidos num prazo máximo de 2 semanas depois da 1ª entrevista.

Além do aspecto de controlo do trabalho de campo do entrevistador, são avaliados os erros de medição e as suas causas: má compreensão do entrevistado ou do entrevistador, falhas do instrumento de notação, etc.



COMPONENTES DA QUALIDADE DO EUROSTAT

2 - Precisão do IE

ERROS NÃO AMOSTRAIS

Erros de Medida

Taxas de Consistência:

[95, 100]- Sexo, Idade, Situação familiar, Relação com o agregado, Procura de trabalho, Reforma, Saber ler, Estar a estudar.

[85, 95] - Condição perante o trabalho, Fonte de rendimento, Situação na profissão, Tipo de contrato.

[70, 85] - Actividade principal, Profissão principal, Horas efectivas, Horas habituais (por intervalos), Nível de instrução.

[50, 70] - Horas habituais, Horas efectivas.



COMPONENTES DA QUALIDADE DO EUROSTAT

2 - Precisão do IE

ERROS NÃO AMOSTRAIS

Erros de Codificação

Referem-se aos erros cometidos quando se atribuem códigos às respostas a perguntas abertas.

O INE realiza alguns testes de avaliação dos erros de codificação:

- Avaliação por recodificação cruzada
- Avaliação por recodificação própria
- Avaliação por comparação com um padrão.

Os erros de codificação detectados pela comparação com um padrão atingem no 3º trimestre de 2002 um máximo de 9% em algumas variáveis.



COMPONENTES DA QUALIDADE DO EUROSTAT

2 - Precisão do IE

ERROS NÃO AMOSTRAIS

Erros de Não-Resposta

Referem-se aos erros cometidos por não se recolher a informação completa em todas as unidades da amostra.

Taxa de não resposta em 2002 - 7% num total de 20 747 U.A

Estes erros são quantificados e corrigidos juntamente com os erros de cobertura através de um método de reponderação chamado calibração.



COMPONENTES DA QUALIDADE DO EUROSTAT

5 - Comparabilidade

A comparabilidade refere-se à medida do impacto das diferenças entre estatísticas, quando comparadas entre áreas geográficas, domínios não geográficos ou períodos de referência

- **Comparabilidade geográfica**

Descrição dos conceitos e métodos que afectam a comparabilidade geográfica dos resultados; assimetrias para os fluxos das estatísticas em espelho;

- **Comparabilidade Temporal**

Descrição da metodologia de tratamento da sazonalidade e da estacionaridade;

Identificação de quebras de série e suas causas;

Descrição dos métodos utilizados no seu tratamento;

- **Comparabilidade entre domínios**

Descrição das diferenças de conceitos e métodos e avaliação das diferenças observadas.



COMPONENTES DA QUALIDADE DO EUROSTAT

6 - Comparabilidade Temporal no IE

O Inquérito ao Emprego tem estado sujeito a várias quebras de série.
Podem distinguir-se essencialmente quatro grandes períodos :

- 1974 a 1982 - Inquérito semestral, só para o Continente.**
(micro dados não recuperáveis)
- 1983 → Inquéritos trimestrais para o País.**
- 1983 a 1992 - Idade activa 12+ anos, desemprego lato.**
- 1993 a 1997 - Idade activa 14+anos, desemprego restrito.**
- 1998 → Idade activa 15+anos, desemprego restrito.**

Bases de amostragem e metodologias diferentes



COMPONENTES DA QUALIDADE DO EUROSTAT

6 - Comparabilidade Temporal no IE

**SÉRIES LONGAS DO
INQUÉRITO AO EMPREGO**

TRATAMENTO DE QUEBRAS

Objectivo do estudo

Corrigir o comportamento aparentemente estranho das séries de estatísticas mais comuns do Inquérito ao Emprego

Actividade, Emprego e Sectores de Actividade

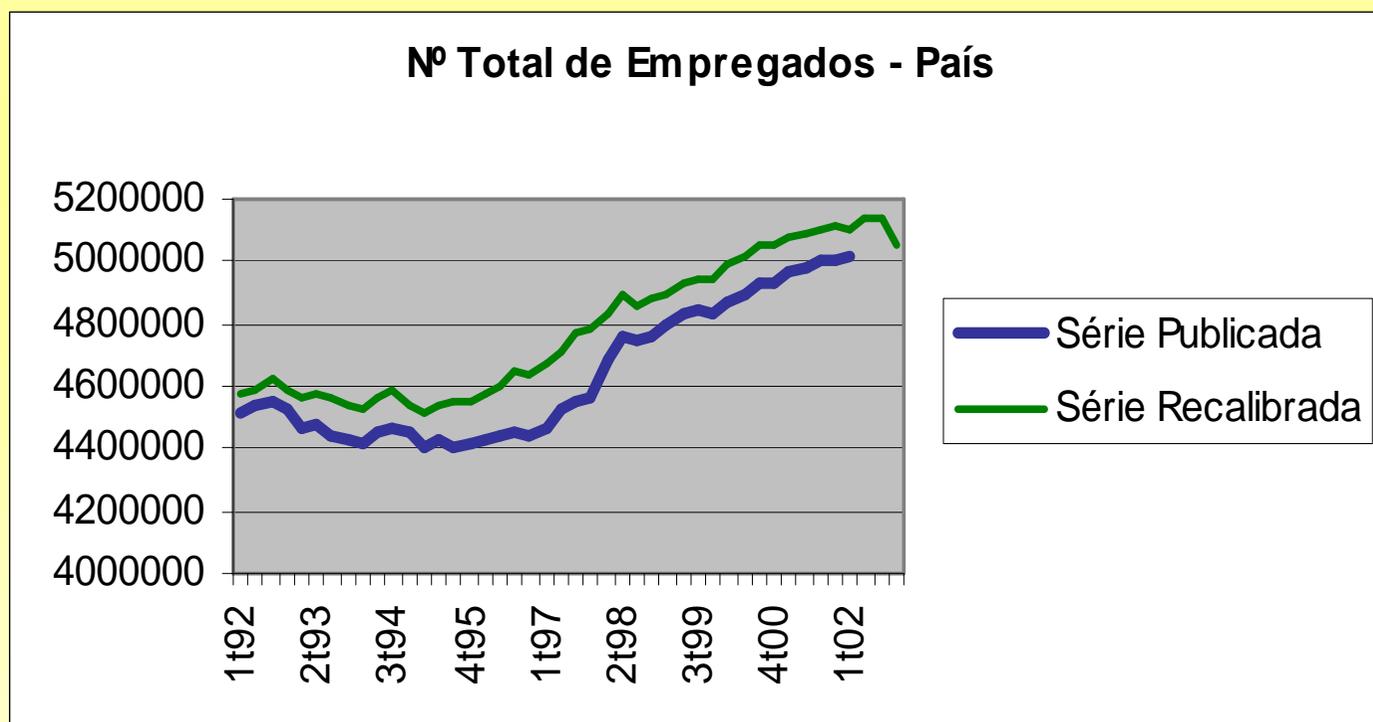
Primeira análise restrita ao período 1992 / 2002

Segunda análise referente ao período 1982/2002

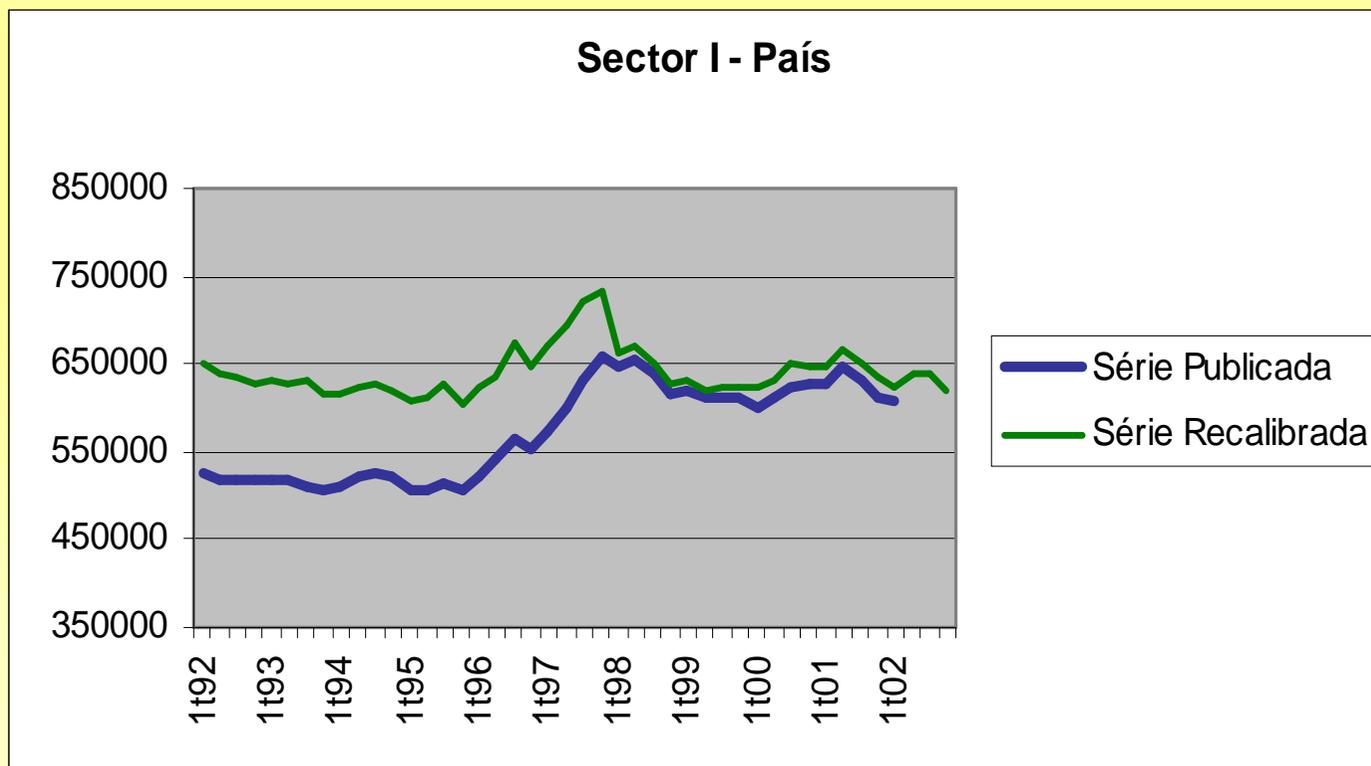
Tópicos

- Qual o aspecto do problema
- O que se consegue através da recalibração fina
- Tratamento estatístico através de séries cronológicas
- Comparação entre o tratamento dos dois períodos
- Objectivos finais

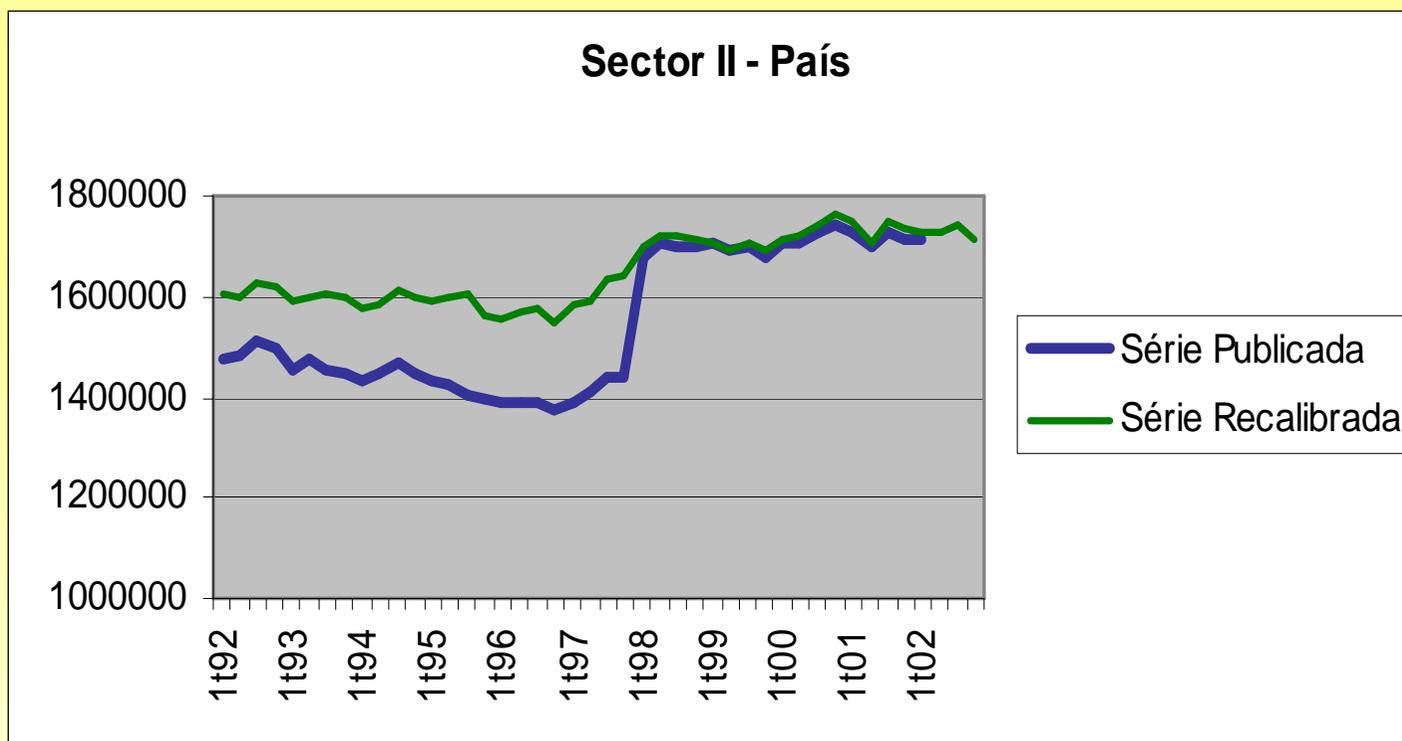
Qual o aspecto do problema ?



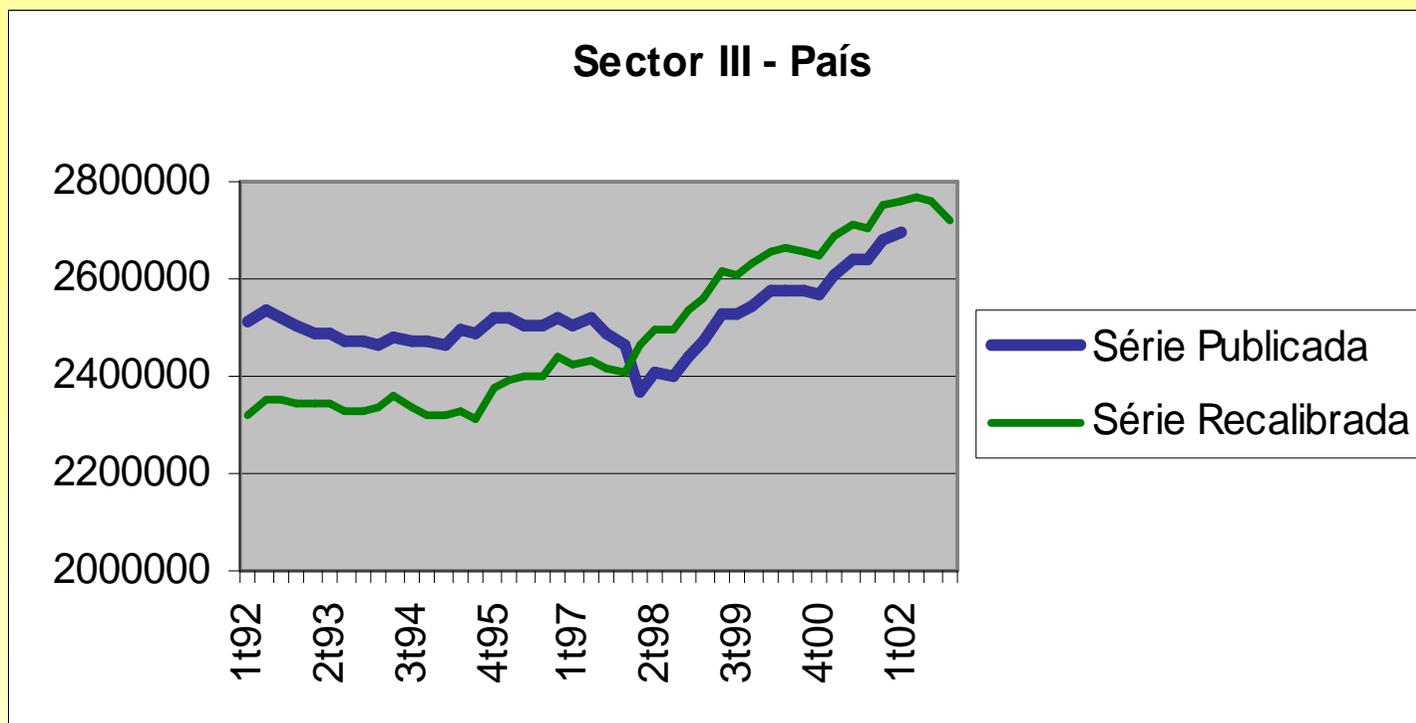
Qual o aspecto do problema ?



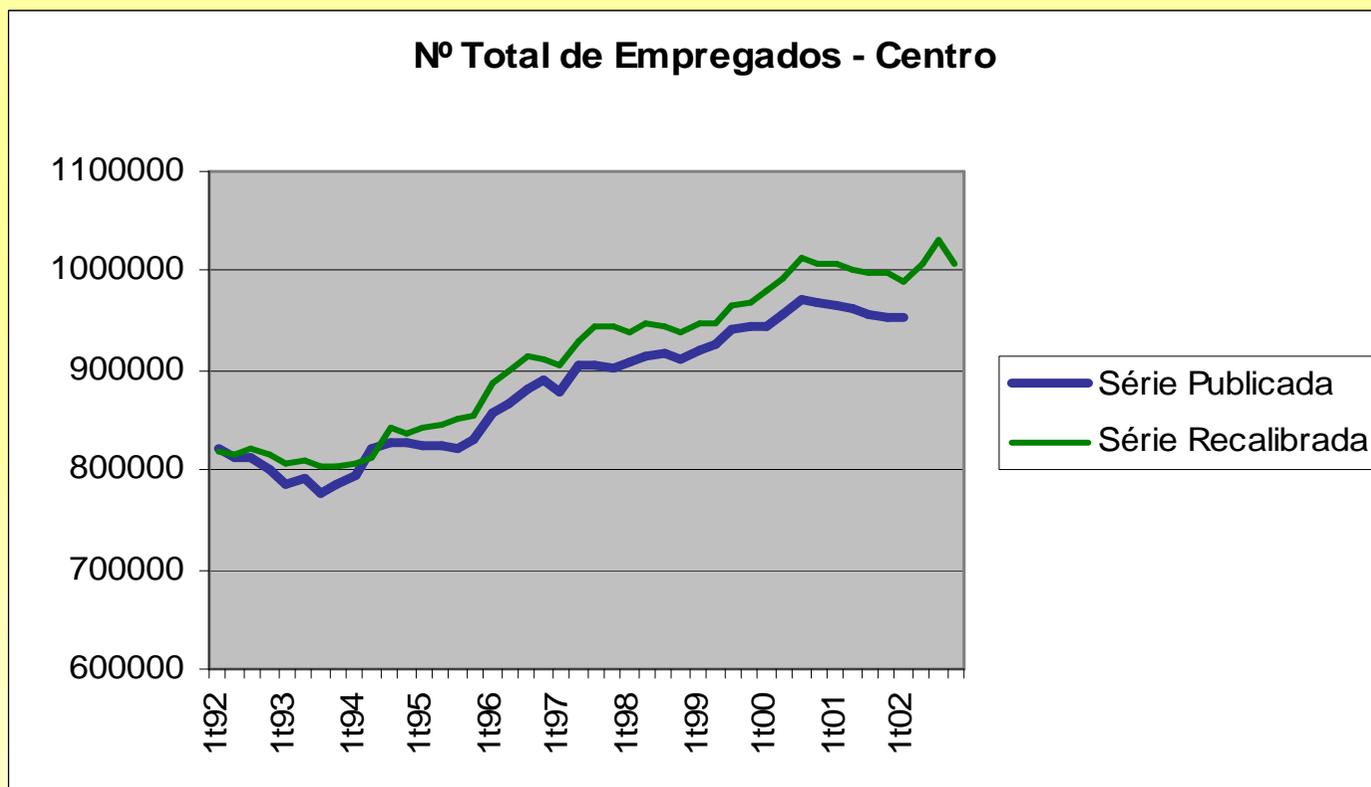
Qual o aspecto do problema ?



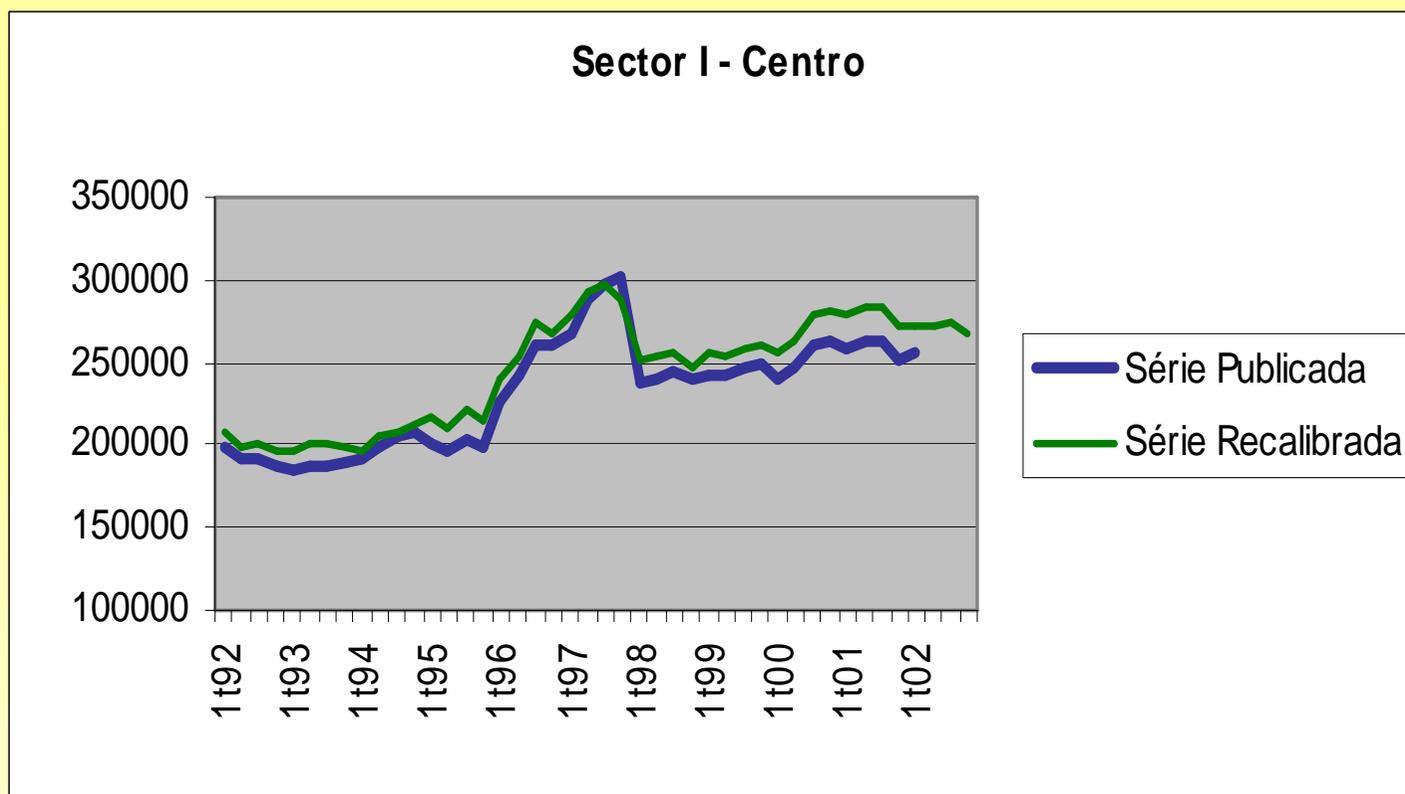
Qual o aspecto do problema ?



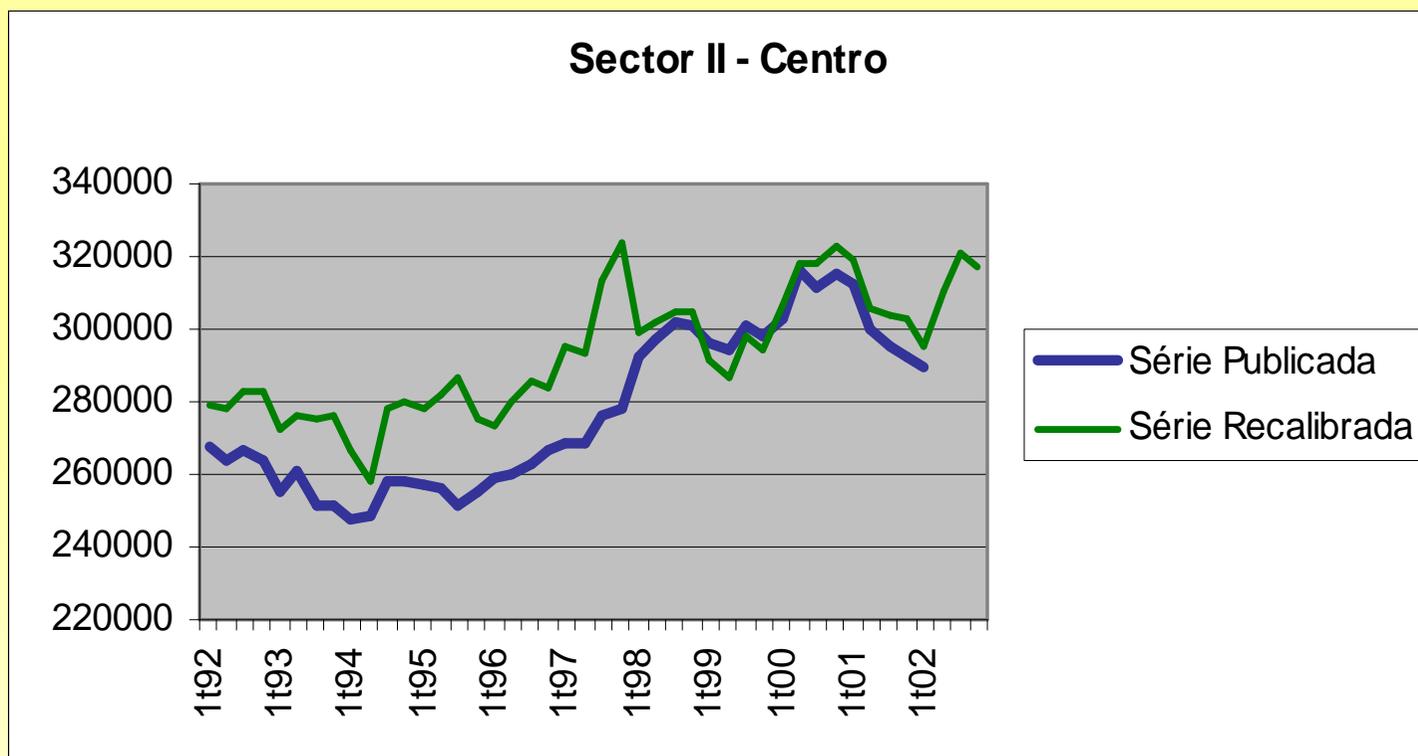
Qual o aspecto do problema ?



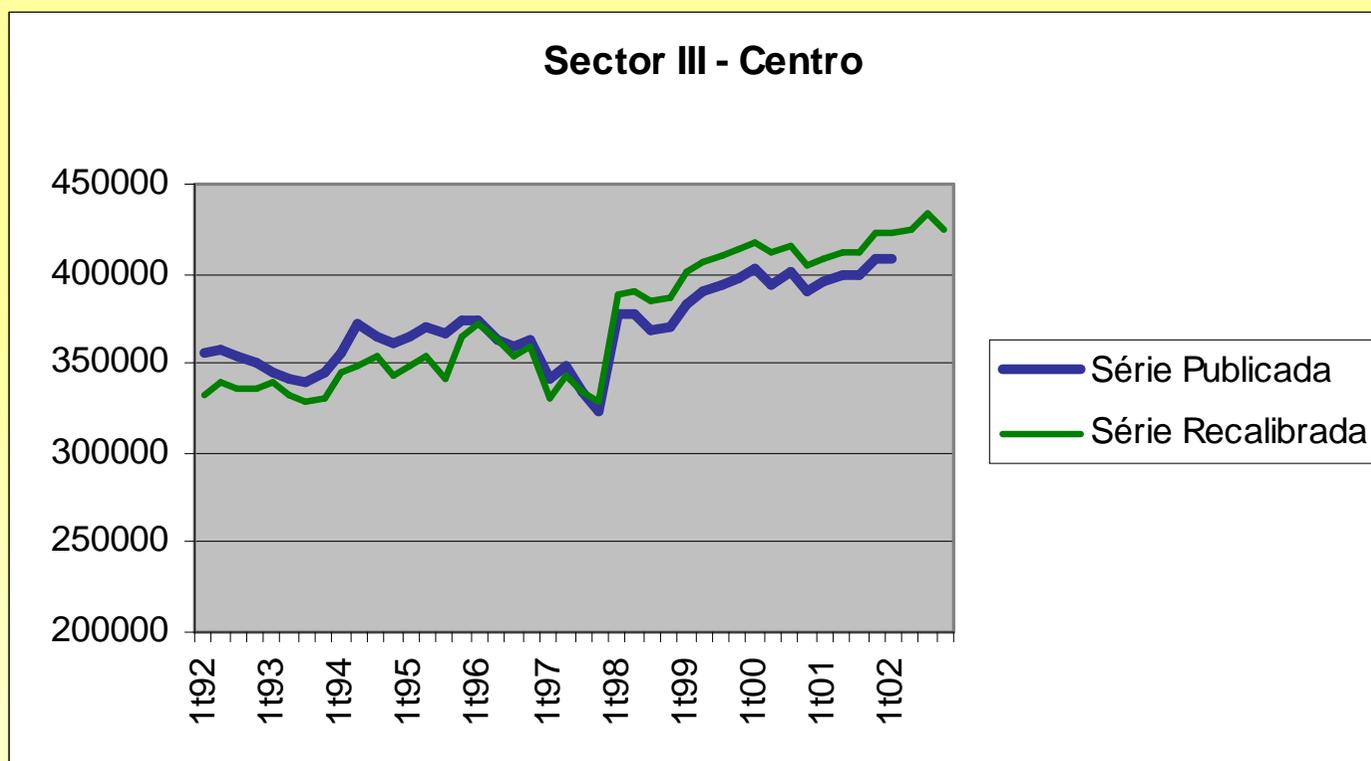
Qual o aspecto do problema ?



Qual o aspecto do problema ?



Qual o aspecto do problema ?





CALIBRAÇÃO

Técnica de pós-estratificação através da qual são atribuídos pesos às unidades observadas por forma a que a soma dos pesos das unidades da amostra pertencentes a uma certa classe reconstitua exactamente o efectivo populacional dessa classe

Estratos utilizados:

- **NUTS II × Género × Classe etária quinquenal**
- **Nuts III × Género**
- **Nuts III × Classes etárias menos desagregadas**
- **Nuts III × Género × Algumas classes etárias importantes**

Efectivos estimados com base nos resultados provisórios dos Censos 2001



Métodos estatísticos de análise de séries cronológicas com intervenção

À parte intervenções que podem ser identificadas no tempo e estimadas, o comportamento da série num certo instante pode ser sempre descrito através de uma mesma função das observações anteriores e dum erro de amplitude controlável

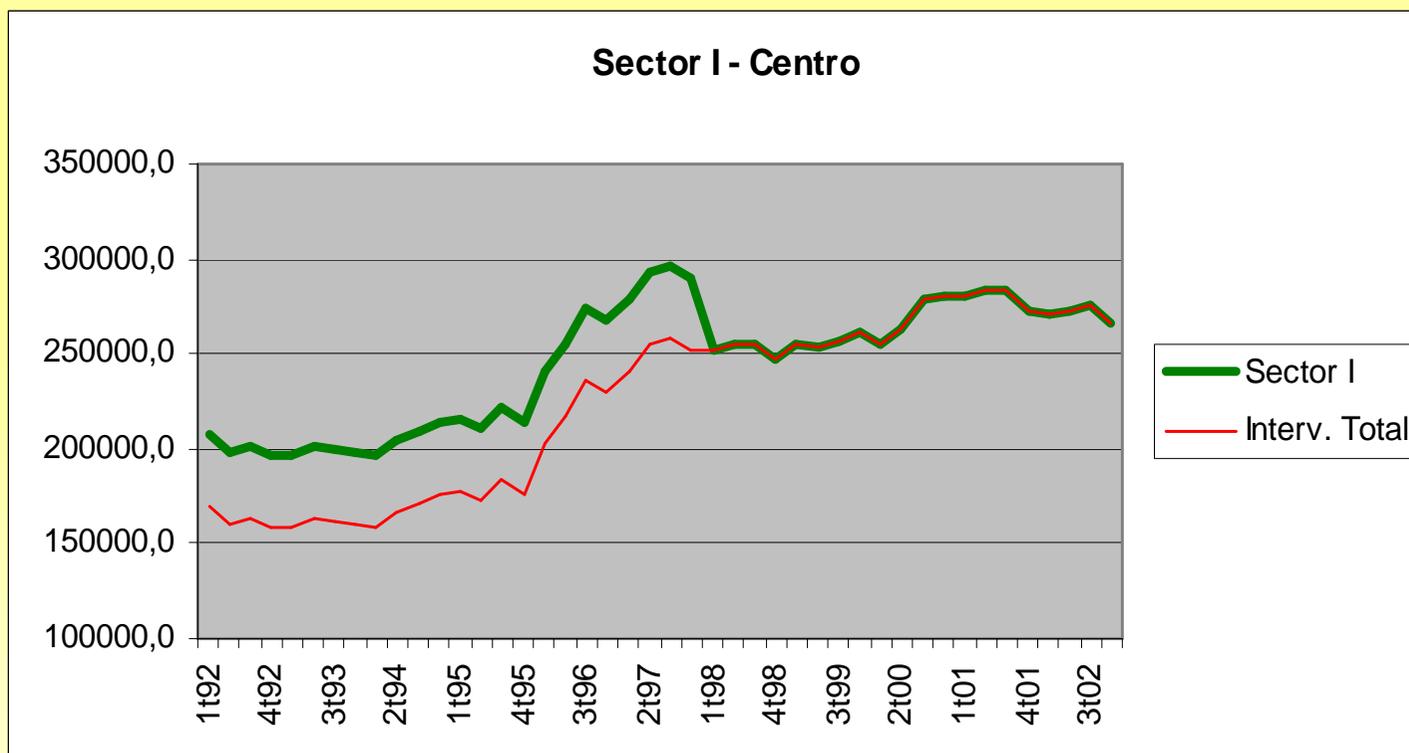
$$Z_t = f(Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots) + \varepsilon_t$$

Uma intervenção representa uma mudança no comportamento da série
Salto, degrau, crescimento ou decrescimento linear, etc

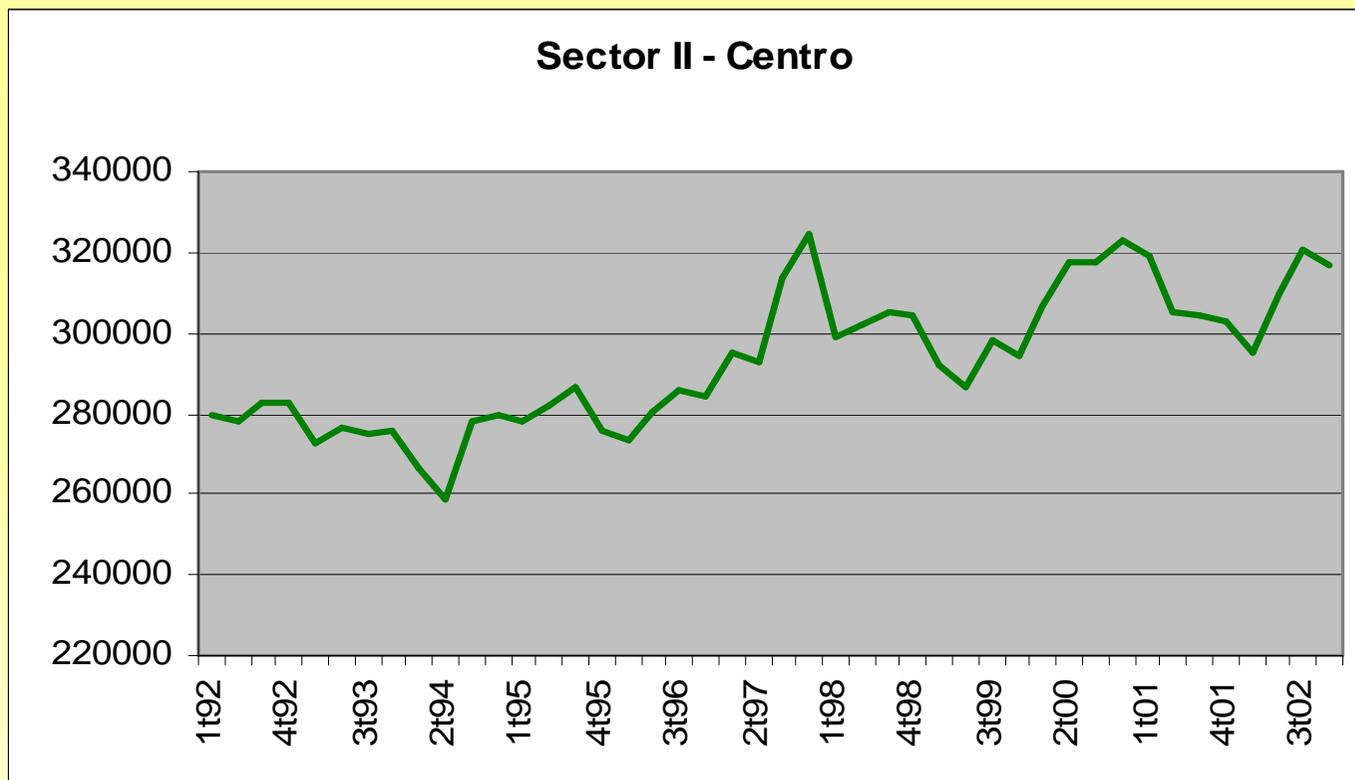
Problemas da análise

- Série pouco longa, especialmente antes da intervenção
- Grande conjunto de séries a tratar simultaneamente
- Séries não independentes com condicionamentos múltiplos
 - Série do emprego é a soma das séries dos sectores de actividade
 - Séries para o país é a soma das correspondentes séries das regiões

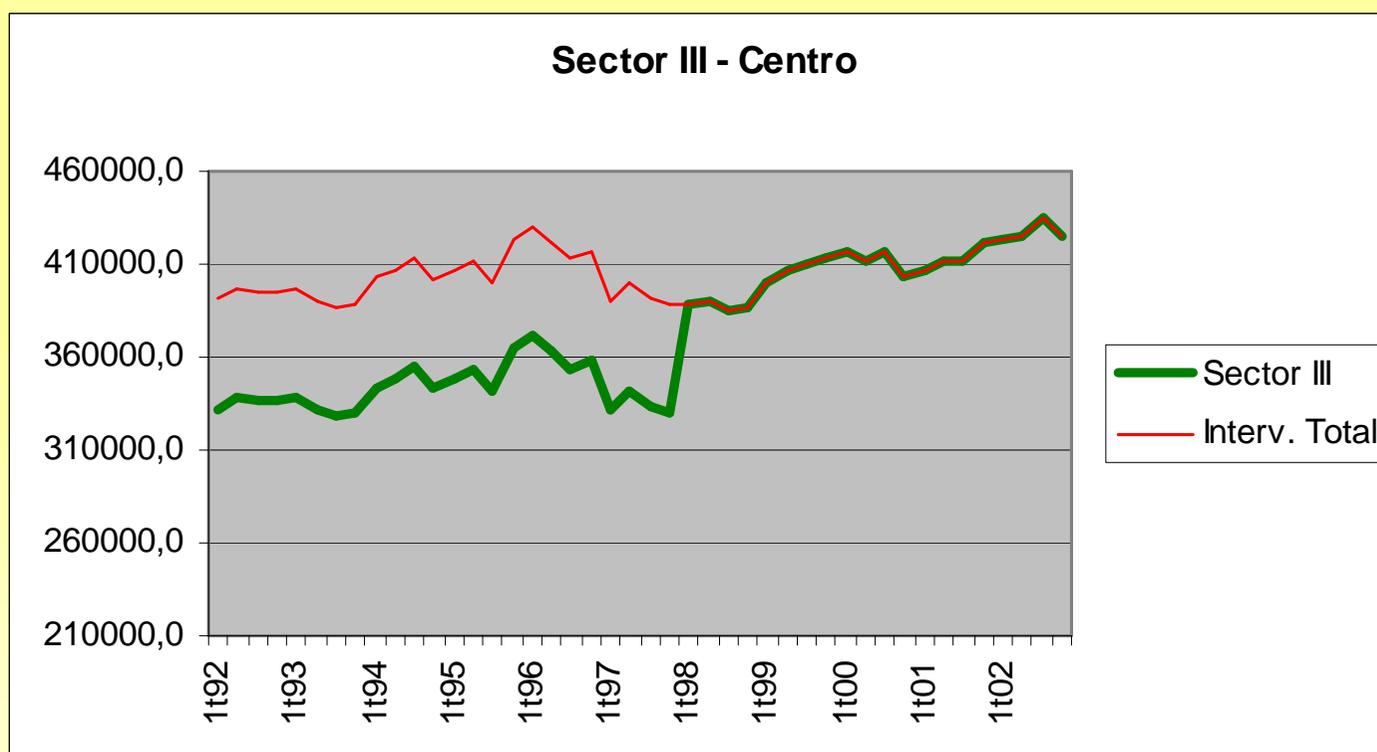
Séries cronológicas com intervenção Método Proporcional com um benchmark



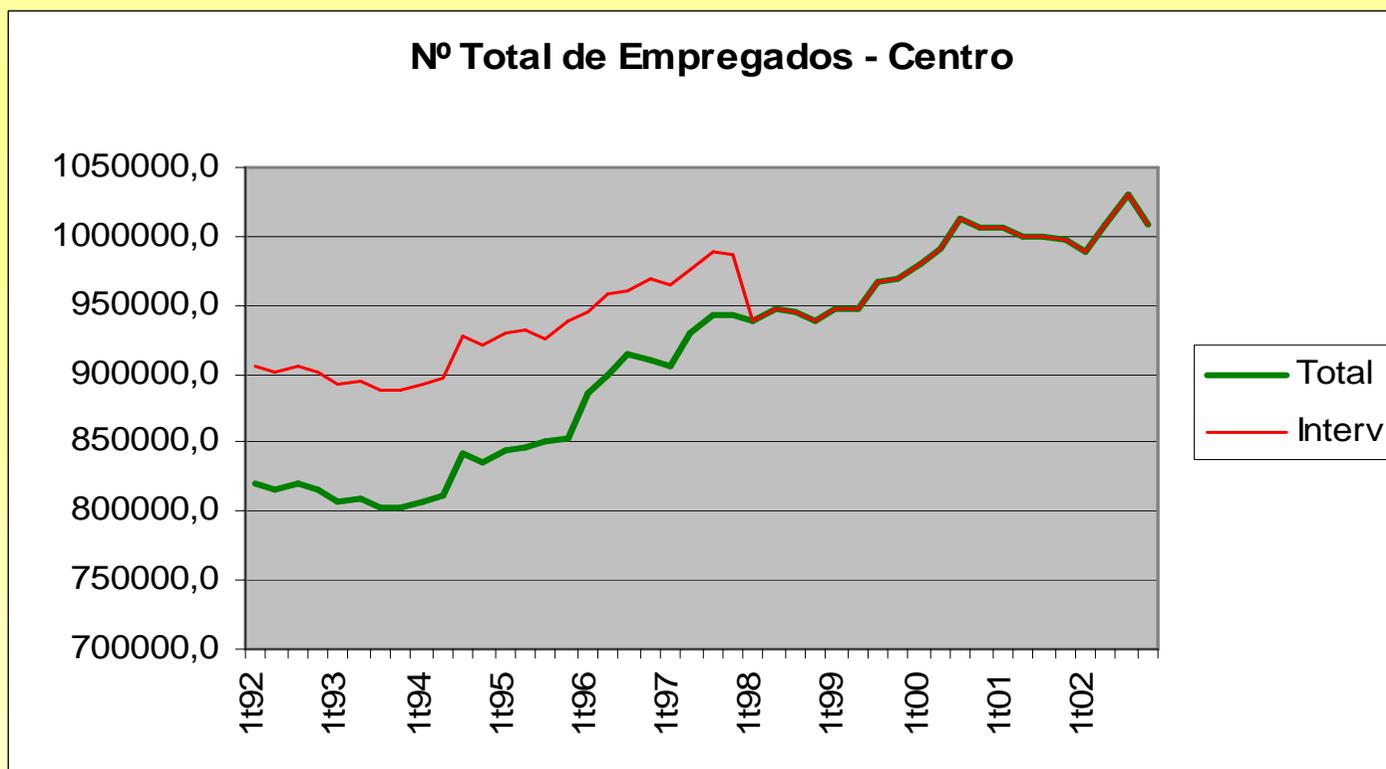
Método Proporcional com um benchmark



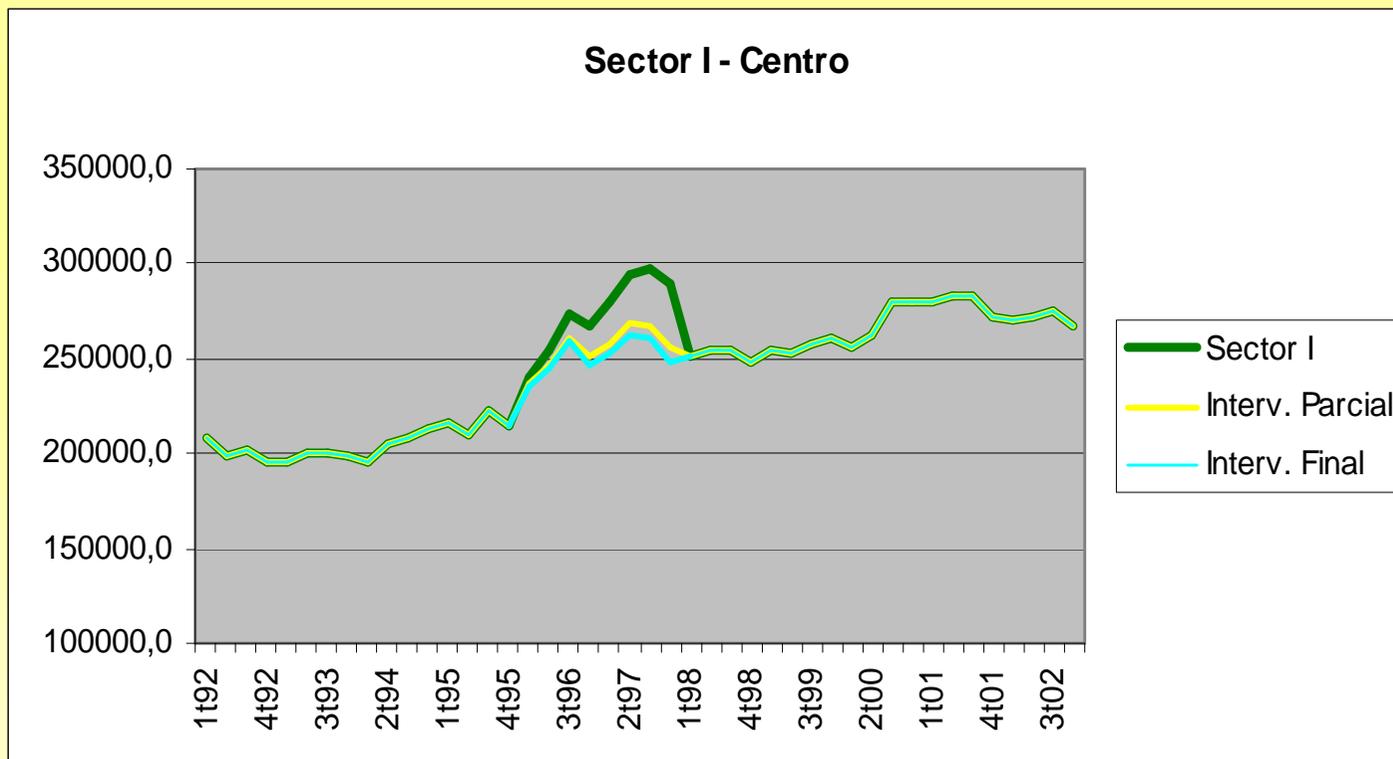
Método Proporcional com um benchmark



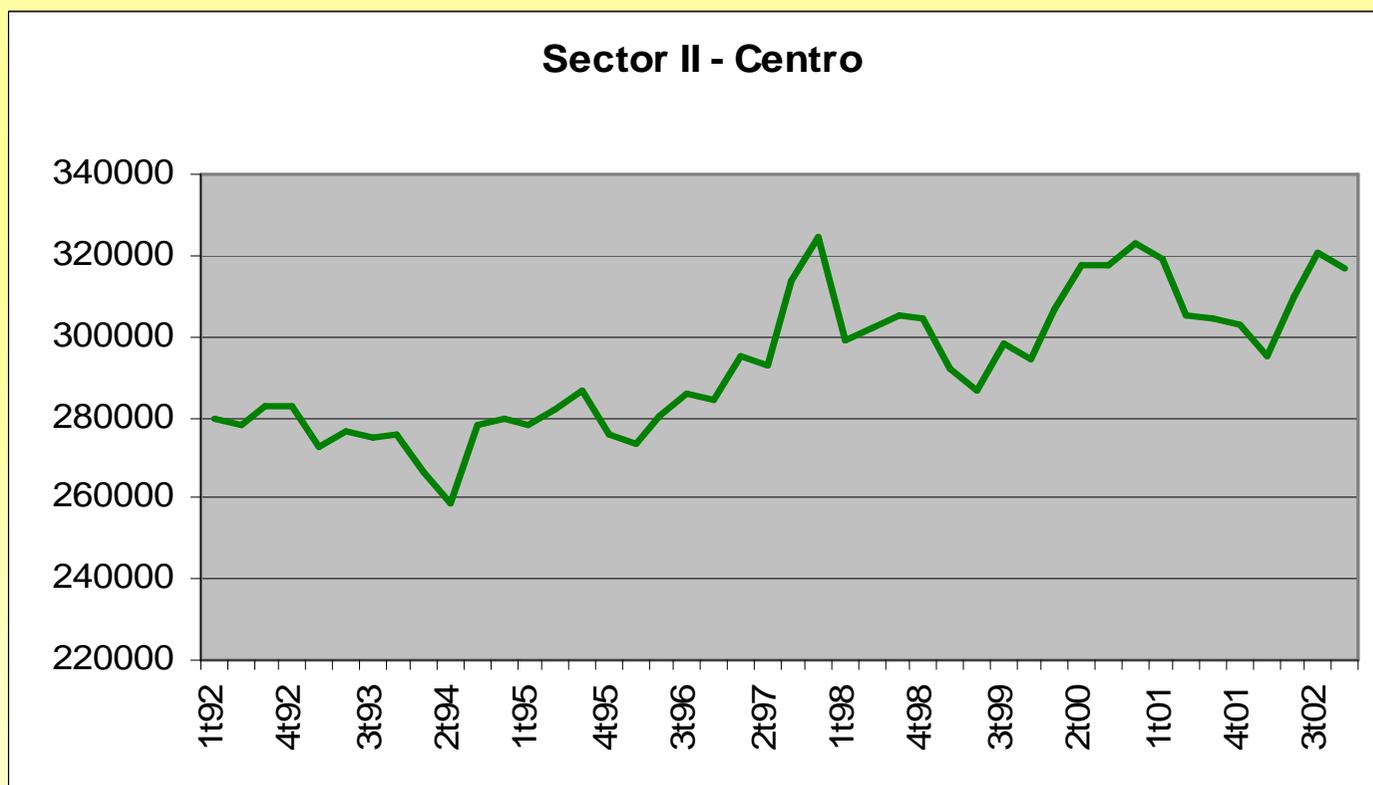
Método Proporcional com um benchmark



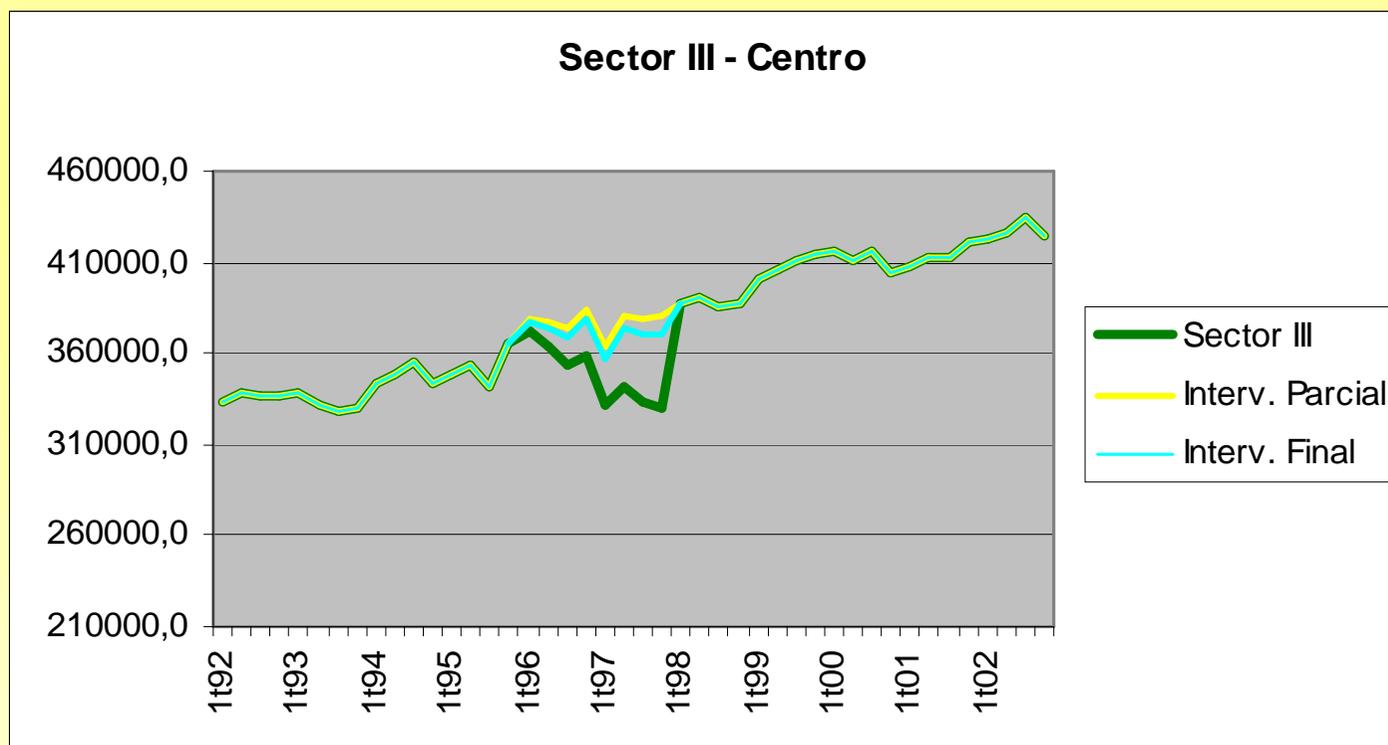
Método proporcional com 2 benchmarks



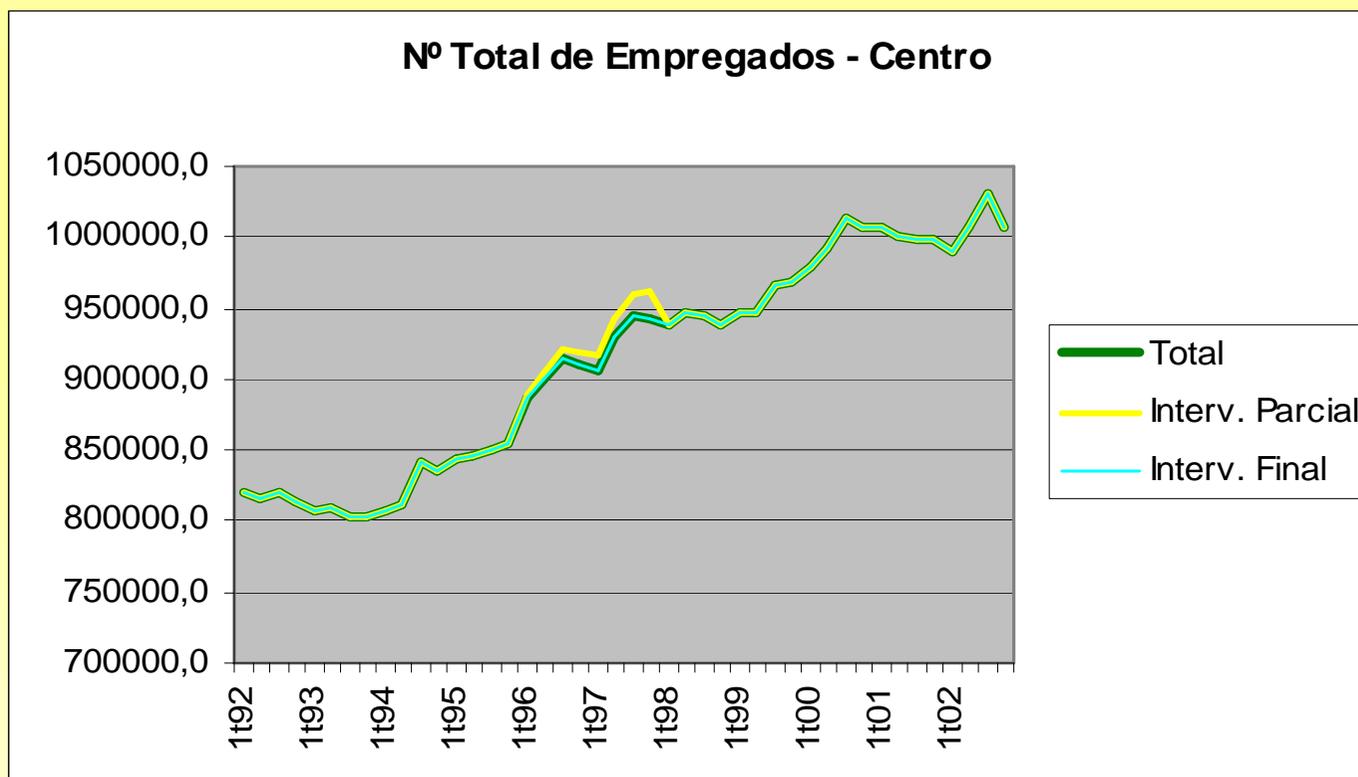
Método proporcional com 2 benchmarks



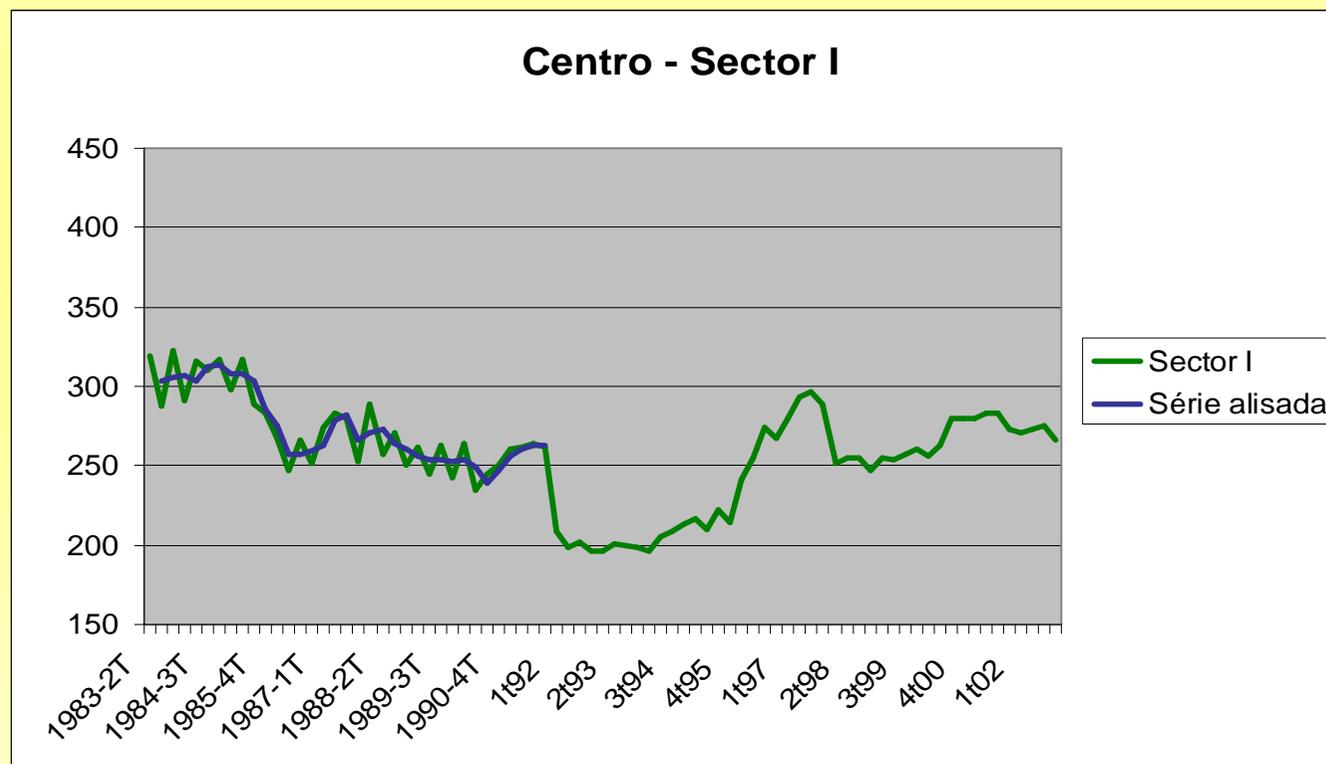
Método proporcional com 2 benchmarks



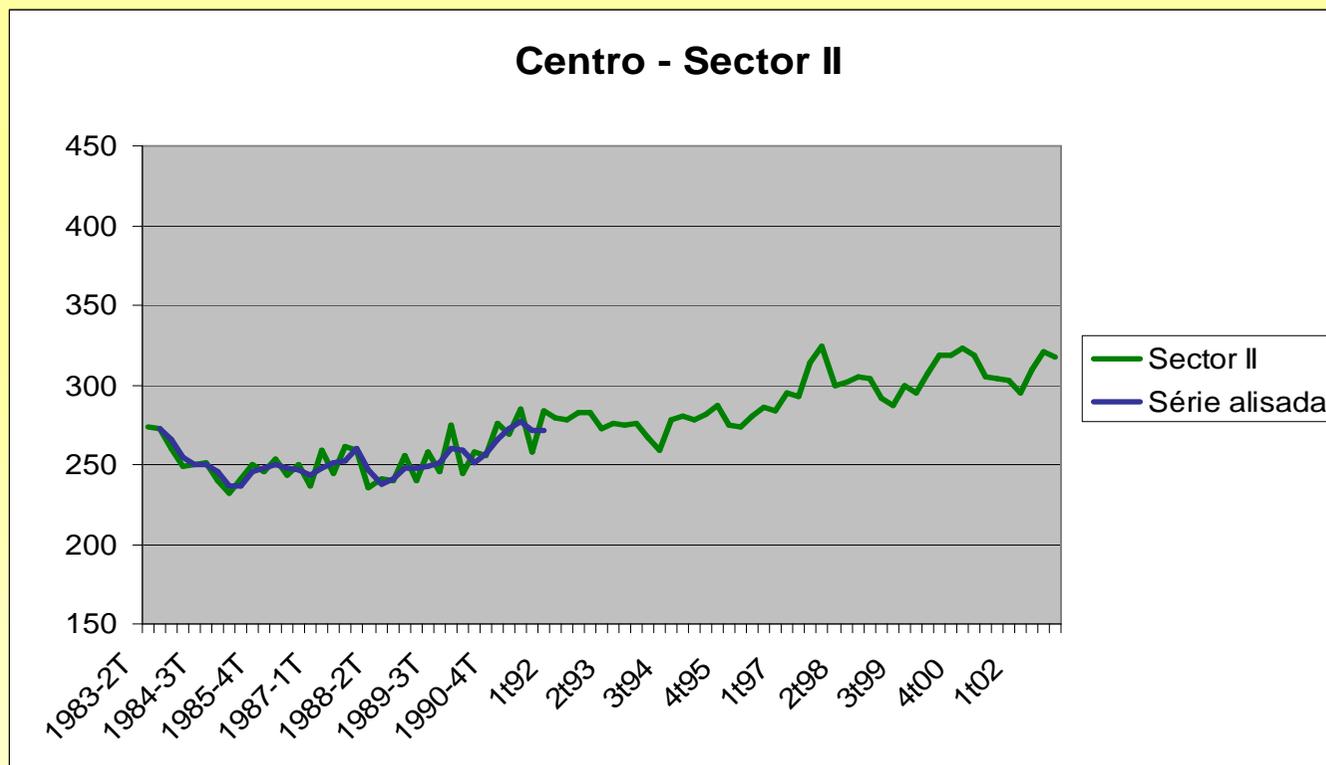
Método proporcional com 2 benchmarks



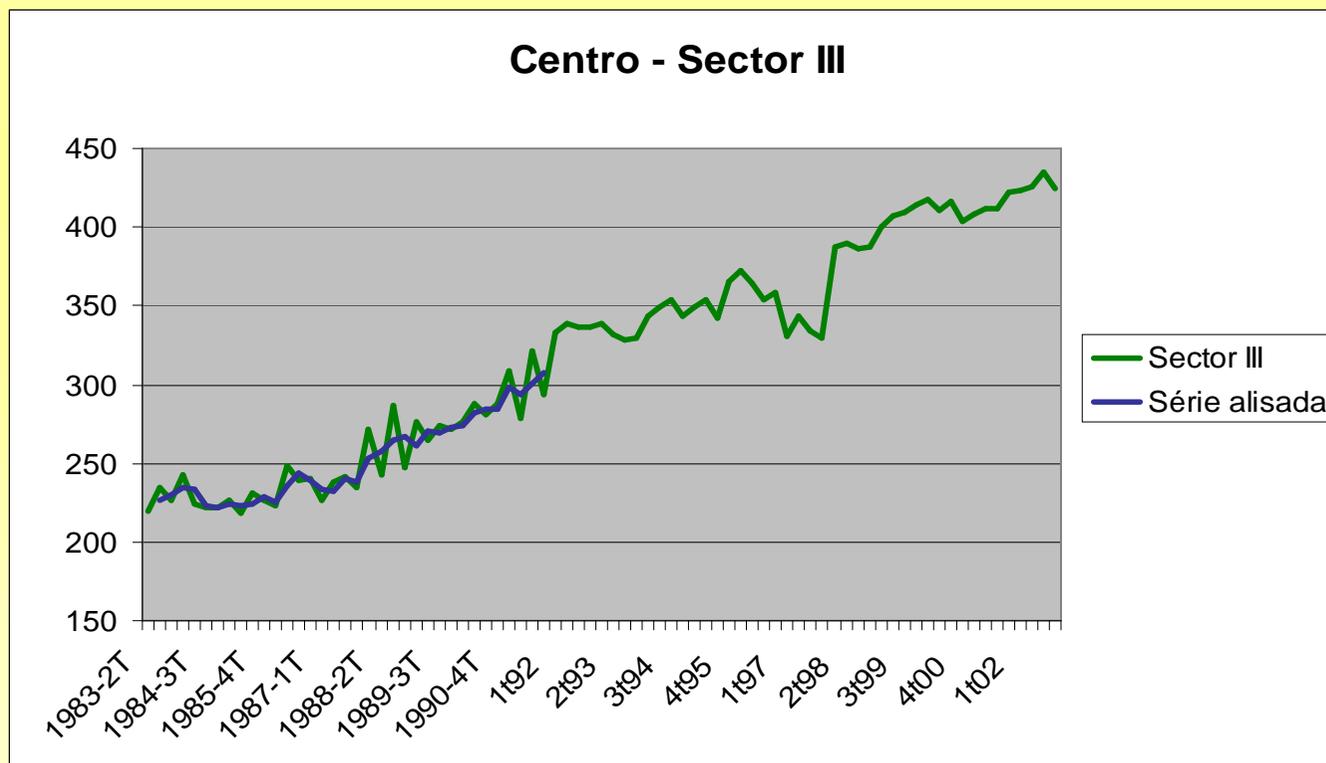
Série completa 1983-2002



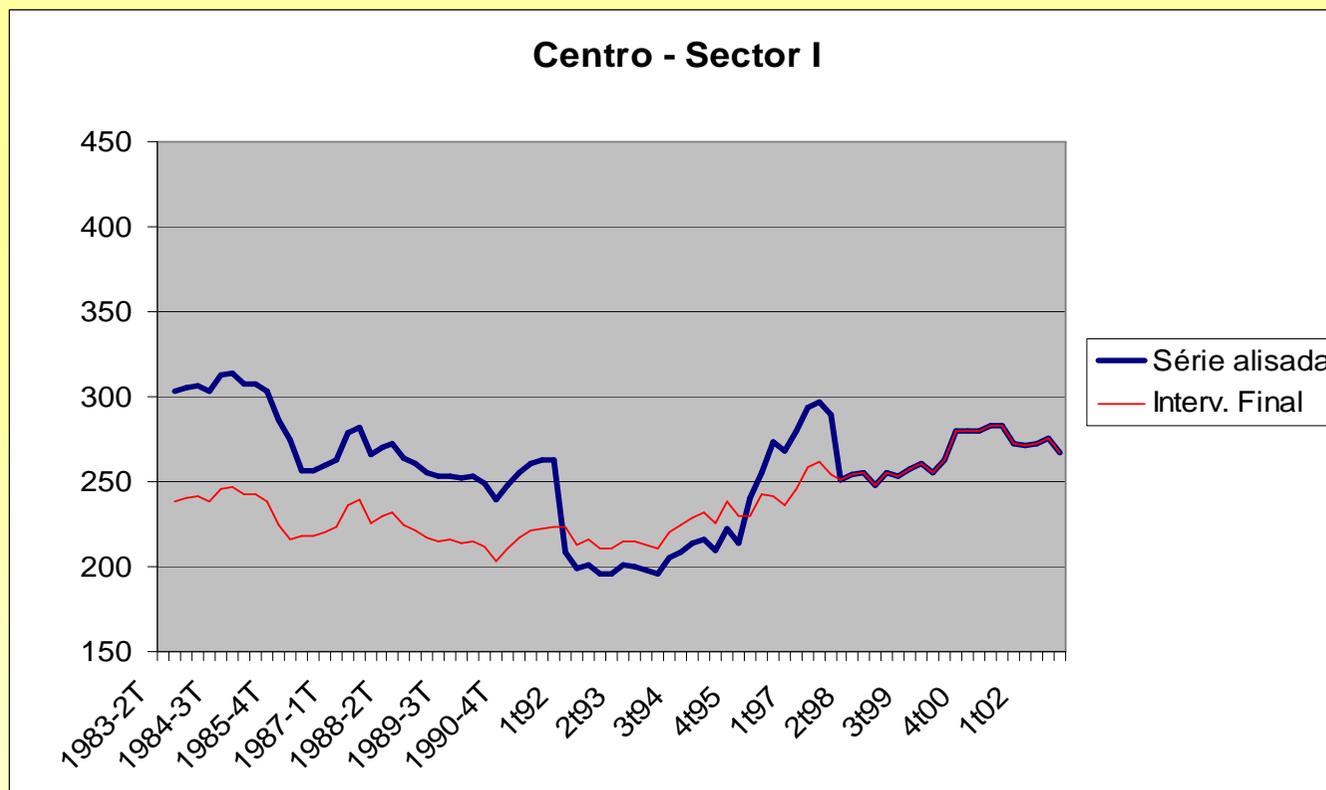
Série completa 1983-2002



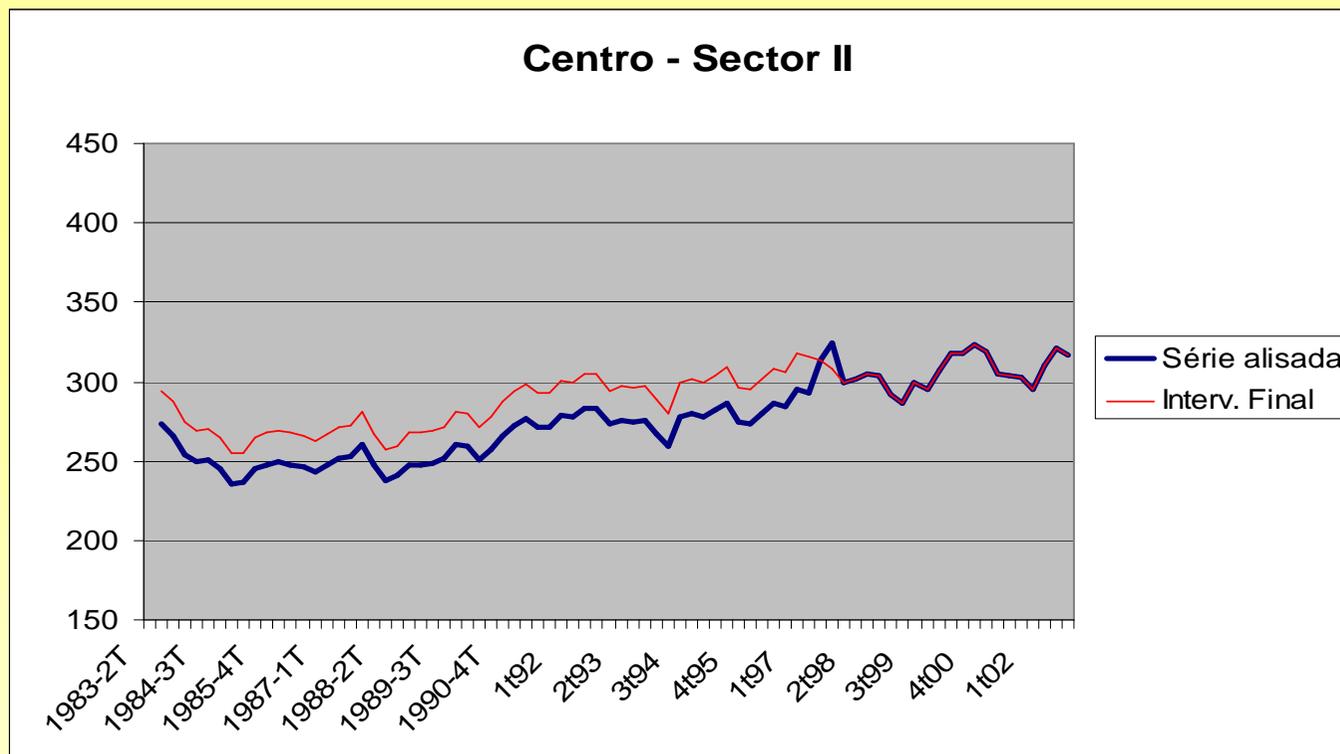
Série completa 1983-2002



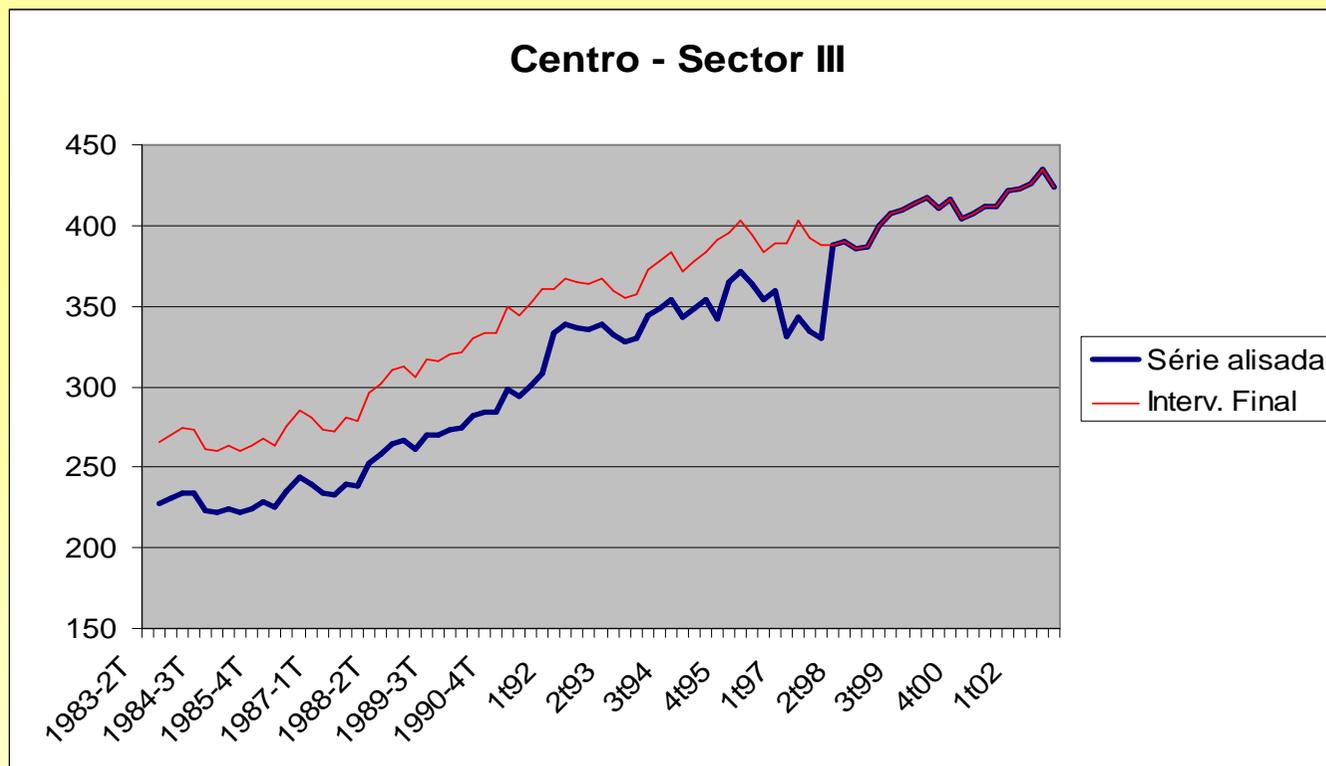
Série completa 1983-2002



Série completa 1983-2002



Série completa 1983-2002



Objectivo Final

Usar as séries corrigidas dos sectores de actividade para calibrar as amostras por NUTS II determinando os pesos unitários que permitem refazer todos os apuramentos de 1983 a 2002.

As séries tratadas para o emprego e sectores de actividade são obtidas como subproduto do método.

O que está feito?

- **Tratamento das séries dos sectores de actividade por NUTS II**
- **Verificação da possibilidade de calibração através das classes de actores de actividade**



COMPONENTES DA QUALIDADE DO EUROSTAT

6 - Comparabilidade Temporal no IE

Transição da Amostra 96 para a Amostra 2001



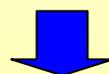
Para permitir uma passagem sem sobressaltos da amostra do IE, criada em 96, e em funcionamento desde o início de 98, para uma nova amostra com base nos Censos 2001, efectuaram-se vários estudos com o fim de escolher uma metodologia de transição adequada, mantendo dentro de limites razoáveis os custos materiais e humanos necessários à realização de toda a operação

Tópicos Metodológicos

- Diferenças dos desenhos amostrais
- Comparação das populações residentes



Diferenças nos Desenhos Amostrais



Focos potenciais de desigualdade entre apuramentos

Amostra de 98

- Amostra-mãe 96
- **Características do Desenho**
 - Independente em cada NUTS II
 - 3 Etapas
(Freguesia, secção, 1 segmento)
 - Dimensão
20747 alojamentos
 - Preparada para apuramentos com
Antiga Geografia

Amostra de 2003

- Amostra-mãe 2001
(mais áreas \Rightarrow mais dispersão)
- **Características do Desenho**
 - Independente em cada NUTS III
 - 2 Etapas (Área, 2 segmentos)
(menos etapas \Rightarrow mais dispersão)
(mais segmentos \Rightarrow mais dispersão)
 - Dimensão
21990 alojamentos
 - Preparada para apuramentos com
Antiga Geografia
Nova Geografia



Comparação das Populações Residentes

Após a escolha da A-M 2001 existiam condições para testar, com base nos resultados provisórios dos Censos 2001, a igualdade das distribuições de características relevantes para as populações residentes nas áreas correspondentes às duas Amostras-Mãe.

Utilizaram-se testes de χ^2 para verificar a igualdade das distribuições relativas a variáveis físicas, tais como

- Género
- Classe etária (várias partições)

e também relativas a variáveis sócio-económicas como

- Alfabetismo
- Frequência e nível de ensino atingido
- Profissão Principal
- Sectores de Actividade
- Situação de emprego/desemprego
- Etc.

tanto a nível nacional como a nível regional



Principais resultados

- Para as variáveis físicas verificou-se uma grande proximidade das distribuições tanto a nível nacional como regional.
- Para as variáveis sócio-económicas embora a proximidade fosse menor existia, mesmo assim, alguma concordância.

Nomeadamente, nas variáveis relacionadas com o emprego por sectores, as diferenças encontradas, ao nível das centenas, não eram estatisticamente significativas, tanto para o país, como para as regiões NUTS II.

CONCLUSÃO

As duas populações em causa são razoavelmente homogéneas



Consequências da análise

Encurtamento de 2 para 1 trimestre do período de sobreposição temporal das duas amostras

Incorporação progressiva da amostra nova na antiga por rotações de $1/6$

Agradecimentos

Estagiárias

Sara Dias; Liliana Ferreira; Sónia Rodrigues

Prof. Kamil Feridun Turkman (FCUL)

Prof. Daniel Muller (ISEG)

ISEGI

BIBLIOGRAFIA

- Documento metodológico do Inquérito ao Emprego
- Relatório Europeu de Qualidade
Inquérito ao Emprego 2001
- Relatório de Estágio de Sara Dias