



LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MESTRADO

MÉTODOS QUANTITATIVOS PARA A DECISÃO ECONÓMICA E EMPRESARIAL

TRABALHO FINAL DE MESTRADO

RELATÓRIO DE ESTÁGIO

**INQUÉRITO ÀS DESPESAS DAS FAMÍLIAS – MÉTODOS DE CORREÇÃO
A *POSTERIORI* RECORRENDO A INFORMAÇÃO AUXILIAR**

DINA SOFIA DOS SANTOS MARQUES

OUTUBRO - 2019



LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MESTRADO

MÉTODOS QUANTITATIVOS PARA A DECISÃO ECONÓMICA E EMPRESARIAL

TRABALHO FINAL DE MESTRADO

RELATÓRIO DE ESTÁGIO

**INQUÉRITO ÀS DESPESAS DAS FAMÍLIAS – MÉTODOS DE CORREÇÃO
A *POSTERIORI* RECORRENDO A INFORMAÇÃO AUXILIAR**

DINA SOFIA DOS SANTOS MARQUES

ORIENTAÇÃO:

**ISABEL PROENÇA
EDUARDA GÓIS**

OUTUBRO - 2019

AGRADECIMENTOS

A realização deste trabalho não seria possível sem o apoio de algumas pessoas.

É um privilégio ter tido pessoas que me são próximas a acompanharem-me ao longo do percurso académico, às quais quero expressar o meu mais sincero agradecimento, em especial ao meu irmão e à minha mãe que sempre apoiaram as minhas escolhas.

Agradeço ao Instituto Nacional de Estatística, pela oportunidade de estágio e à Dra. Eduarda Góis pela orientação e disponibilidade demonstrada ao longo do trabalho. Uma palavra de agradecimento a toda a equipa pelo apoio e conhecimento transmitido ao longo deste processo.

Quero também agradecer à coordenação do mestrado por toda a disponibilidade. Um agradecimento especial à minha orientadora, professora Doutora Isabel Proença, por todo o incentivo e dedicação ao longo da realização do relatório de estágio.

RESUMO

O presente relatório final de mestrado tem por base o estágio curricular realizado no Instituto Nacional de Estatística (INE). O objetivo do estágio prende-se com o estudo do Inquérito às Despesas das Famílias (IDEF) e a melhoria da informação disponibilizada pelo mesmo através da criação de métodos de correção *a posteriori* recorrendo a informação auxiliar.

É pretendida uma aproximação dos resultados apresentados pelo IDEF aos resultados de outras fontes externas de informação. A calibração é elaborada no IDEF para variáveis sociodemográficas, a fim de utilizar também uma variável económica, incorporou-se informação disponível pelas Contas Nacionais (CN). Esta análise surgiu da necessidade de testar métodos alternativos de calibração de forma a mitigar os erros característicos de inquéritos amostrais sobre estudos sociais, onde são inquiridos os indivíduos.

A informação das CN utilizada consiste no consumo privado (despesa de consumo final) das famílias no território económico por função (COICOP) a preços correntes (anual). Para a realização do objetivo foi utilizado o *software* SAS e a macro CALJACK que permite a realização da calibração para diversas variáveis quantitativas e qualitativas.

Conclui-se que as CN são uma boa fonte de informação auxiliar para a calibração do IDEF uma vez que foi possível identificar um aumento da qualidade e precisão das observações das variáveis após calibração. A estrutura da informação disponibilizada pelo IDEF aproximou-se da estrutura das CN.

ABSTRACT

The present master final work is based in a curricular internship done at *Instituto Nacional de Estatística* (INE). This internship had the objective of studying an survey, “Inquérito às Despesas das Famílias” (IDEF) and, at the same time, improve its information, through the creation of methods that make corrections *a posteriori*, using auxiliary information.

This work pretends to make an approximation of the results presented by IDEF to the ones presented by external entities. The calibration is executed on IDEF for sociodemographic variables and, in order to use an economic variable, some information about National Accounts was incorporated. This analysis was necessary to test alternative calibration methods, to mitigate the errors that are a characteristic of surveys about societal issues where individuals are inquired.

The information used from National Accounts is the private consumption (spendings with final consumption) made by families in the economic field, by function (COICOP) and with current prices (annual). In order to achieve the objective, the software SAS and the macro CALJACK were used, allowing the calibration for several variables, both quantitative and qualitative.

In sum, National Accounts are a good source of auxiliary information for the calibration of IDEF, since it was possible to identify an improvement in the quality and precision of its estimates. The information structure provided by IDEF got close to the CN structure.

ÍNDICE

AGRADECIMENTOS	I
RESUMO	II
ABSTRACT	III
LISTA DE FIGURAS	IV
LISTA DE TABELAS	V
SIGLAS/ACRÓNIMOS	VI
1. INTRODUÇÃO	1
2. INSTITUTO NACIONAL DE ESTATÍSTICA	2
3. INQUÉRITO ÀS DESPESAS DAS FAMÍLIAS (IDEF)	3
3.1. MÉTODO DE AMOSTRAGEM	6
3.2. MÉTODO DE RECOLHA	8
4. MÉTODOS DE CORREÇÃO A POSTERIORI RECORRENDO A INFORMAÇÃO AUXILIAR	9
4.1. CALIBRAÇÃO	9
4.2. CONTAS NACIONAIS	11
4.3. REVISÃO BIBLIOGRÁFICA	12
5. CALIBRAÇÃO ATRAVÉS DA MACRO CALJACK	17
6. APRESENTAÇÃO E ANÁLISE DE RESULTADOS	23
7. CONCLUSÃO	27
REFERÊNCIAS	29
ANEXOS	31
ANEXO 1 – COICOP: CLASSIFICAÇÃO INDIVIDUAL DO CONSUMO DOS ADP POR FUNÇÃO	31
ANEXO 2 – CALIBRAÇÃO NO SAS	34
ANEXO 3 – VARIÁVEIS MACRO	38
ANEXO 4 – CALIBRAÇÃO PARA CADA REGIÃO DA NUTS_II.	40
ANEXO 5 – COEFICIENTE DE VARIAÇÃO	46

LISTA DE FIGURAS

Figura 1 – Comandos utilizados para a Construção do ficheiro “novo” e das novas variáveis.....	35
Figura 2 – Comandos usados para a criação do ficheiro “idef15”.....	36
Figura 3 – Criação de um novo ficheiro para cada região das NUTS_II.	36
Figura 4 – Macro CALJACK: Margens	37
Figura 5 – Macro CALJACK: Norte	41
Figura 6 – Macro CALJACK: Centro	41
Figura 7 – Criação das variáveis <i>des16</i> , <i>des17</i> e <i>des18</i> (Centro).....	42
Figura 8 – Macro CALJACK: Lisboa	42
Figura 9 – Macro CALJACK: Alentejo	43
Figura 10 – Criação da variável <i>des26</i> (Alentejo)	43
Figura 11 – Macro CALJACK: Algarve.....	44
Figura 12 – Macro CALJACK: Região Autónoma dos Açores.....	44
Figura 13 – Criação da variável <i>des27</i> (Açores)	45
Figura 14 – Macro CALJACK: Região Autónoma da Madeira.....	45
Figura 15 – Criação das variáveis <i>des13</i> , <i>des14</i> , <i>des23</i> , <i>des25</i> (Madeira). 46	

LISTA DE TABELAS

Tabela 1 – Comparação entre os totais das variáveis COICOP para diferentes fontes de informação.....	21
Tabela 2 – Estrutura da despesa total anual média por ADP em 2015, por divisão da COICOP segundo as Contas Nacionais, em percentagem.....	23
Tabela 3 – Estrutura da despesa total anual média por ADP, realizada em território nacional, por divisões da COICOP, NUTS_II, em 2015/16 (calibração sem informação auxiliar das CN) em percentagem.....	24
Tabela 4 – Estrutura da despesa total anual média por agregado, realizada no território nacional, por classe da COICOP, NUTS_II, em 2015/16 (calibração com informação auxiliar das CN), em percentagem.....	25
Tabela 5 – Níveis da COICOP	31
Tabela 6 - Comparação do Coeficiente de Variação para a região Norte...	46
Tabela 7 - Comparação do Coeficiente de Variação para o Centro.....	47
Tabela 8 - Comparação do Coeficiente de Variação para Lisboa	47
Tabela 9 - Comparação do Coeficiente de Variação para o Alentejo.....	48
Tabela 10 - Comparação do Coeficiente de Variação para o Algarve	48
Tabela 11 - Comparação do Coeficiente de Variação para a Região Autónoma dos Açores	48
Tabela 12 - Comparação do Coeficiente de Variação para a Região Autónoma da Madeira	49

SIGLAS/ACRÓNIMOS

- ADP - Agregado Doméstico Privado
- *CAPI – Computer Assisted Personal Interview*
- CN – Contas Nacionais
- COICOP – Classification of Individual Consumption by Purpose
- EU – União Europeia
- EUA – Estados Unidos da América
- EUROSTAT – Serviço de Estatística da União Europeia
- FNA – Ficheiro Nacional de Alojamentos
- GRID - células retangulares de 1x1 Km
- *HBS – Household Budget Surveys*
- IDEF – Inquérito às Despesas das Famílias
- INE – Instituto Nacional de Estatística
- *INSEE - Institut National de la Statistique et des Études Économiques*
- IOF – Inquérito aos Orçamentos Familiares
- IRDF – Inquérito às Receitas e Despesas Familiares
- *NIH – National Institutes of Health*
- NUTS - Nomenclatura das Unidades Territoriais para Fins Estatísticos
- *PSU – Primary Sampling Unit* (em português, Unidade Primária da Amostra)
- *SAS - Statistical Analysis System*
- SCN – Sistema de Contas Nacionais
- SEC – Sistema Europeu de Contas Nacionais e Regionais
- SEE – Sistema Estatístico Europeu
- SEN – Sistema Estatístico Nacional

1. INTRODUÇÃO

O mundo que nos rodeia é mais facilmente compreendido se for quantificado. A Estatística fornece informação essencial para entender a cidadania, as atividades empresariais e o funcionamento da economia, sendo um instrumento fundamental na construção de políticas económicas e sociais. O Instituto Nacional de Estatística (INE) é responsável pela produção e divulgação de estatísticas oficiais nacionais de qualidade e relevantes a toda a sociedade.

O INE realiza periodicamente o Inquérito às Despesas das Famílias (IDEF) que fornece informação estatística relativamente à estrutura das despesas, distribuição dos rendimentos e níveis de conforto dos agregados domésticos privados residentes em Portugal. No IDEF são usados procedimentos de controlo automático da qualidade durante a fase de recolha de dados. Para além disso, os dados recolhidos são ponderados de forma a serem representativos da população residente através de um método de calibração que utiliza variáveis auxiliares que asseguram o ajustamento para os totais conhecidos da população. Tradicionalmente estas variáveis auxiliares são variáveis demográficas.

Ao longo do trabalho pretende-se testar métodos alternativos de calibração de resultados, usando informação de fontes externas, de modo a corrigir os efeitos resultantes da heterogeneidade das taxas de resposta e das diferenças na qualidade das respostas revelados pelos distintos estratos populacionais. Assim, o objetivo deste trabalho trata do desenvolvimento métodos de correção *a posteriori* recorrendo a informação auxiliar.

A utilização de informação disponível sobre a população para melhorar a qualidade das estimativas pode ocorrer na fase da conceção do método de amostragem ou na fase de calibração. A informação auxiliar quando utilizada na fase de estimação, fase essa naturalmente posterior à recolha de dados, tem como objetivo obter estimativas das observações das variáveis de interesse mais próximas dos valores da população. A calibração tem como base a utilização de uma ou mais variáveis auxiliares que não foram tidas em consideração no momento da conceção do método de amostragem. “Ao constatar-se, *a posteriori*, que a amostra não reproduz fielmente as estatísticas conhecidas para as

variáveis auxiliares, incorpora-se, então, a informação auxiliar no processo de estimação, isto é, as variáveis vão ser explicitamente incluídas nas fórmulas dos estimadores” (Coelho, Pereira, Pinheiro e Xufre, 2017). As Contas Nacionais (CN) Portuguesa produzidas e divulgadas pelo INE publicam, anualmente, informação acerca do consumo privado (despesas de consumo final) o que permite a utilização da mesma como informação auxiliar para a calibração do IDEF. O objetivo é analisar a qualidade da informação disponibilizada pelas CN com fim à sua utilização no cálculo das estimativas do IDEF.

O trabalho está dividido em cinco grandes pontos. Inicialmente é apresentada a instituição de acolhimento para a elaboração do estágio, o INE, seguindo-se a exposição do que é o IDEF, do seu método de amostragem e do método de recolha elaborado para o mesmo. No ponto 4 são abordados os métodos de correção *a posteriori* recorrendo a informação auxiliar e a revisão bibliográfica acerca desse tema. Num subponto é exposto o que são as Contas Nacionais e o porquê da sua utilização enquanto informação auxiliar para a calibração do IDEF. Neste ponto insere-se ainda a exploração acerca dos erros que ocorrem ao longo do processo de sondagem e que, conseqüentemente, originam uma diminuição da qualidade das variáveis calibradas. No ponto 5 é descrito o programa utilizado, o SAS, e a macro CALJACK, entrando assim nos procedimentos utilizados para a realização da calibração e conseqüentemente para a construção dos estimadores. Por fim são apresentados e analisados os resultados que foram obtidos com a elaboração do projeto.

2. INSTITUTO NACIONAL DE ESTATÍSTICA

Garantir padrões elevados de qualidade e rigor aliados à divulgação de informação atualizada numa sociedade em constante mudança é fundamental para que a informação estatística seja relevante aos seus utilizadores.

O Instituto Nacional de Estatística (INE) é o organismo central de produção e divulgação de estatísticas oficiais em Portugal e a sua atividade é enquadrada por um quadro jurídico próprio, nacional e europeu. Tem como missão, “produzir e divulgar de forma eficaz, eficiente e isenta, informação estatística oficial de qualidade, relevante para toda a Sociedade” (INE, 2014). O INE é um instituto público integrado na Administração Indireta do Estado, com personalidade

jurídica e autonomia administrativa, inserido na estrutura da Presidência do Conselho de Ministros

As estatísticas são produzidas com independência técnica, sem entrar em incumprimento com as normas provindas do Sistema Estatístico Nacional (SEN) e do Sistema Estatístico Europeu (SEE). O SEE é uma parceria entre o Eurostat (autoridade estatística da União Europeia) e as autoridades estatísticas dos Estados Membros. As autoridades estatísticas são constituídas pelos institutos nacionais de estatística e outras autoridades públicas nacionais às quais são delegadas competências para a produção de estatísticas oficiais. Em regime de delegação de competências por parte do conselho diretivo do INE, cabe ao mesmo a responsabilidade de garantir a coordenação da metodologia, a qualidade e a divulgação dessa informação estatística.

O INE encontra-se sediado em Lisboa com delegações no Porto, Coimbra, Évora e Faro. O Serviço Regional de Estatística dos Açores e a Direção Regional de Estatística da Madeira, na produção das estatísticas de âmbito nacional, encontram-se equiparados às delegações do INE em termos da Lei do SEN.

As estatísticas são disponibilizadas a toda a sociedade, de forma a assegurar o cumprimento do segredo estatístico e sendo o seu acesso tendencialmente gratuito. As operações estatísticas respeitam critérios de qualidade, impostos a nível nacional e internacional, de onde se salienta a precisão e atualidade, onde a atividade estatística se encontra em permanente evolução, leva à ocorrência de revisões de resultados, um processo natural estabelecido através da Política de Revisões. A informação estatística é um instrumento indispensável para o desenvolvimento económico e social na sociedade atual.

3. INQUÉRITO ÀS DESPESAS DAS FAMÍLIAS (IDEF)

Nas economias de mercado, os cidadãos adquirem o seu rendimento através da venda da sua própria força de trabalho, do retorno pelo investimento e das transferências do Estado e outras entidades, designadamente privadas e/ou do exterior. O mercado recompensa, assim, a iniciativa, o esforço e os riscos assumidos pelos indivíduos. As diferenças geradas na distribuição do rendimento originam desigualdade no poder de compra entre os cidadãos.

Disponer de informação detalhada acerca da riqueza e do rendimento das famílias e das suas opções de poupança e despesa, por características sociodemográficas constitui um elemento precioso para a política económica.

É neste contexto que se enquadra o Inquérito às Despesas das Famílias (IDEF), realizado pelo INE de cinco em cinco anos, e que constitui a principal fonte estatística sobre a distribuição das despesas das famílias em Portugal. O inquérito é realizado com o objetivo de recolher informação sobre as despesas monetárias e não monetárias de cada agregado doméstico privado (ADP), sendo que em simultâneo recolhe dados de enquadramento sobre as características sociodemográficas e sobre o rendimento dos seus membros individualmente e como um todo. As despesas não monetárias correspondem ao autoconsumo (produção própria), à autolocação (autoavaliação de um valor hipotético do alojamento para ADP proprietários ou usufrutuários gratuitos do mesmo), ao autoabastecimento (bens ou serviços obtidos de estabelecimento explorado pelo agregado) e aos recebimentos em géneros e salários em espécie.

O IDEF é uma das operações estatísticas mais consolidadas do INE e com uma utilização crescente dos vários setores e entidades nacionais e internacionais. Pertence a uma série estatística, designada Orçamentos Familiares que teve início em 1967/68 com a realização do Inquérito às Receitas e Despesas das Famílias (IRDF). De meados dos anos 70 e até 2000 designou-se Inquérito aos Orçamentos Familiares (IOF). A terminologia do Inquérito às Despesas das Famílias (IDEF) surgiu em 2005/06 e mantém-se até à atualidade. No global foram realizados nove inquéritos: 1967/68, 1973/74, 1980/81, 1989/90, 1994/95, 2000, 2005/06, 2010/11 e 2015/16.

O IDEF é realizado com uma periodicidade quinquenal, principalmente de forma a caracterizar os comportamentos de despesa dos agregados familiares, assegurando a disponibilização de informação necessária ao Índice de Preços do Consumidor e ao Consumo Final apurado no âmbito das Contas Nacionais Portuguesas, e fornecer dados sobre as quantidades alimentares necessárias à atualização periódica da Balança Alimentar Portuguesa. Os resultados do inquérito são ainda apropriados pelo projeto europeu *Household Budget Survey (HBS)*, que procede à compilação das recolhas de dados sobre orçamentos

familiares em todos os países da União Europeia. Ao nível nacional, o IDEF permite ainda a obtenção de indicadores de conforto da população residente através da recolha de dados sobre condições de habitabilidade e alguns bens disponíveis nos lares portugueses. Ao aliar o IDEF a outros inquéritos às famílias podem ser elaborados diversos estudos do binómio rendimento-despesa. Assim, a análise do IDEF é importante para a constituição de políticas económicas e sociais de forma célere por parte do Estado.

Vários Estados Membros realizam inquéritos às despesas das famílias, o que permitiu, através de um acordo entre os mesmos, a criação do *Household Budget Survey (HBS)* para a comparação das condições de vida na União Europeia (Office for Official Publications of the European Communities, 2003).

Assim, o Eurostat, em conjunto com os institutos nacionais de estatística dos países que concordaram com a participação no HBS, elaborou recomendações de forma a processar e publicar informação estatística comparável estabelecendo conceitos, métodos, estruturas e padrões técnicos. Esses inquéritos são realizados por iniciativa dos países, sendo a recolha dos dados executada pelos Institutos de Estatística Nacionais. Cada país possui, para o seu inquérito respetivo, liberdade para decidir os objetivos, metodologia, programação e atribuição de recursos. O aumento da coerência na elaboração do inquérito tem vantagens evidentes, porém o Eurostat considera que a regularidade e a periodicidade de realização do inquérito são aspetos mais relevantes.

O agregado doméstico privado (ADP) é o conjunto de indivíduos que partilha o mesmo alojamento e o mesmo orçamento para a provisão de bens essenciais à vida, independentemente da existência ou não de laços familiares. Os empregados domésticos que coabitam no alojamento são considerados membros do ADP.

Para cada agregado doméstico privado é definido um indivíduo de referência, cujas características pessoais são usadas na classificação do agregado. O critério adotado em Portugal e recomendado pelo Eurostat classifica como indivíduo de referência a pessoa que mais contribui para os rendimentos do agregado doméstico privado.

3.1. MÉTODO DE AMOSTRAGEM

A amostragem tem como principal objetivo a obtenção de informação necessária no momento certo e ao menor custo possível. A elaboração de cada inquérito é determinada por circunstâncias específicas, constrangimentos e objetivos de cada país.

O IDEF apresenta como população-alvo todos os agregados domésticos privados com residência principal no território nacional no período de referência da recolha dos dados (*INE – Documento Metodológico*). A base de amostragem utilizada é constituída pelos alojamentos familiares excluindo alojamentos coletivos (lares, hospitais, prisões, etc.). A unidade estatística de observação tanto pode ser o ADP como o indivíduo. A população excluída, em Portugal e na maioria dos Estados Membros, não excede os 1% da população total e é mais significativa junto de grupos particulares, como idosos.

Na maioria dos Estados Membros, os inquéritos HBS, em que se integra o IDEF, utilizam a amostragem estratificada e multietápica, que consiste em estabelecer uma hierarquia de unidades de amostragem. Na primeira etapa é escolhida uma amostra aleatória de unidades primárias (*PSU*), na segunda retira-se, de cada unidade primária da amostra, uma amostra aleatória de unidades secundárias e assim consecutivamente, podendo o processo ter o número de etapas desejado.

A amostra, utilizada pelo IDEF, caracteriza-se por ser probabilística estratificada multietápica (seleção de conglomerados na primeira etapa e de alojamentos na segunda etapa) e transversal. A Nomenclatura das Unidades Territoriais para Fins Estatísticos (NUTS) é o sistema hierárquico que, à semelhança dos restantes Estados membros, divide o território Português em regiões. A seleção da amostra do IDEF foi dimensionada segundo o nível II da NUTS, constituído por sete regiões: Norte; Centro; Área Metropolitana de Lisboa; Alentejo; Algarve; Região Autónoma dos Açores; e Região Autónoma da Madeira. Essas regiões são tidas em consideração na distribuição das unidades primárias e secundárias aquando da realização do método de amostragem.

Na definição da amostra do IDEF, as *PSU* são constituídas geograficamente por uma ou mais células contíguas da GRID INSPIRE2¹ de 1km². A posição de cada um dos edifícios dos Censos 2011 na GRID foi determinada através da sobreposição espacial com as células GRID, onde se conclui que apenas 43% das células têm alojamentos de residência principal. As células GRID foram desenvolvidas pelo Eurostat para o território Europeu e são selecionadas de forma sistemática e com probabilidade proporcional à dimensão do número de alojamentos de residência principal. As células GRID permitem assim a concentração das unidades a serem inquiridas em áreas possibilitando a redução dos custos de transporte e de outros custos de recolha dos dados. Para além disso, permite um melhor acompanhamento e controlo, fundamental para uma melhoria da taxa de resposta e qualidade dos dados recolhidos. As *PSU* foram selecionadas sistematicamente com probabilidade proporcional à dimensão do número de alojamentos de residência principal para cada NUTS_II.

Os alojamentos são as unidades secundárias, os quais foram selecionados de forma sistemática dentro das *PSU*. Assim dentro das unidades secundárias são inquiridos todos os agregados domésticos privados e indivíduos que aí tenham a sua residência principal. A utilização de alojamentos como unidades secundárias possui vantagens dado que são geralmente mais estáveis, facilmente identificáveis e providenciam uma cobertura mais completa da população.

Para a construção da amostra é usada informação recolhida através dos Censos. O Recenseamento da População e da Habitação ou Censos é realizado a cada dez anos. A base de amostragem utilizada para o IDEF 2015/2016 foi extraída do Ficheiro Nacional de Alojamentos (FNA), elaborado em 2013 a partir dos Censos 2011. No que respeita às bases de amostragem dos inquéritos às famílias, é particularmente relevante a atualização do FNA através de informação auxiliar interna, designadamente a que resulta das operações estatísticas do SEN. No FNA os alojamentos possuem variáveis de identificação e localização (códigos, coordenadas e morada), variáveis de caracterização (situação do

¹ GRID oficial desenvolvida pelo Eurostat para o território Europeu

alojamento e área útil) e variáveis de controlo (disponibilidade do alojamento, data e fonte de atualização).

A amostra deve ser representativa da população geograficamente e ao longo do tempo, de forma a minimizar os efeitos sazonais. O período de inquirição do IDEF é de um ano, sendo que cada ADP é objeto de recolha de dados durante 14 dias sendo, assim, distribuídos pelas 26 quinzenas. Essa distribuição da amostra final em termos de unidades de alojamento é realizada também ao nível das NUTS_II.

3.2. MÉTODO DE RECOLHA

De forma a inquirir os agregados domésticos privados são realizadas diversas entrevistas (*INE - Documento Metodológico*). As primeiras visitas têm como objetivo a explicação do preenchimento das cadernetas de consumo e a recolha de dados relativos ao alojamento, ao agregado e aos indivíduos através de *Computer Assisted Personal Interview (CAPI)* (Módulo I). Nesta primeira entrevista é inquirida informação relativa ao alojamento, agregado, indivíduos, educação, conforto e bens de equipamento e receitas monetárias líquidas. De seguida, pretende-se o acompanhamento do autopreenchimento das cadernetas de consumo quinzenal do ADP e do indivíduo que são fornecidas pelo entrevistador, em papel, permanecendo com o indivíduo durante os 14 dias de recolha (Módulos II e III). A caderneta de consumos diários do agregado destina-se ao registo das despesas correntes do ADP e para cada um dos seus indivíduos (Módulo II). A caderneta individual é utilizada para indivíduos do agregado maiores de idade e que apresentem a necessidade de preencher individualmente algumas despesas (Módulo III). As cadernetas devem refletir todo o consumo realizado pelo ADP durante esse período de tempo, em bens e serviços, identificando o modo de consumo, compra, autoconsumo, autoabastecimento recebimentos gratuitos ou recebimentos a título de salário/complemento. Por fim existe a recolha de informação numa perspetiva retroativa quanto a consumos mensais, trimestrais e anuais por *CAPI* (Módulo IV). Os dados relativos ao consumo quinzenal, mensal e trimestral são depois anualizados. O período de referência dos dados corresponde ao momento da entrevista relativamente às variáveis de caracterização do alojamento, agregado

e indivíduo, conforto e bens de equipamento. Para as receitas monetárias líquidas do indivíduo o período de referência é o ano fiscal anterior ao ano de início dos inquéritos.

Para a caracterização do consumo é utilizada a classificação do consumo individual por objetivo (COICOP) que é uma classificação definida pelas Nações Unidas, adaptada pelo Eurostat para a aplicação no SEE. A harmonização das classificações e do sistema de codificação das variáveis essenciais é fundamental para uma análise comparativa ao nível da UE. A COICOP classifica os bens e serviços em doze categorias principais de despesas de consumo (bens alimentares e bebidas não alcoólicas; bebidas alcoólicas, tabaco e narcóticos; vestuário e calçado; habitação, água, eletricidade, gás e outros combustíveis; mobiliário, artigos de decoração, equipamento doméstico e manutenção corrente da habitação; saúde; transportes; comunicações; lazer, recreação e cultura; educação; restaurantes e hotéis; bens e serviços diversos). A divisão das despesas segundo a COICOP permite examinar como as famílias distribuem os seus gastos e monitorizar os efeitos causados pelas alterações dos preços ao longo do tempo. No primeiro nível da COICOP a mesma divide-se em 12 classes, designadas Divisões. Cada Divisão é subdividida para o segundo nível da COICOP em subclasses designadas de Grupos. Os Grupos, por sua vez desdobram-se num terceiro nível, designado Classe (veja-se anexo 1 – tabela 5).

4. MÉTODOS DE CORREÇÃO A POSTERIORI RECORRENDO A INFORMAÇÃO AUXILIAR

4.1. CALIBRAÇÃO

A informação disponível sobre a população-alvo deve ser empregue na conceção do método de amostragem, antes da recolha dos dados, de forma a tornar o inquérito mais eficiente. Dispor de informação prévia acerca da população permite a realização de uma correta estratificação. Um desenho da amostra mais eficiente leva a um aumento da qualidade dos dados observados. Contudo essa informação auxiliar pode apenas estar presente na fase de calibração, ou seja, numa fase posterior à recolha dos dados. Nesse caso, a informação auxiliar é usada para a obtenção de estimativas dos valores das

variáveis mais próximas dos valores da população, quando comparadas com as estimativas realizadas com os ponderadores clássicos subjacentes ao desenho da amostra. A calibração é assim um método utilizado *a posteriori*, onde, na construção de estimadores, os valores observados são calibrados de modo a serem corrigidos os efeitos resultantes da heterogeneidade das taxas de resposta e das diferenças na qualidade de resposta. Os métodos de calibração têm por base uma ou mais variáveis auxiliares, para as quais se conhece, no caso de variáveis auxiliares quantitativas, a média ou o total na população, ou no caso de variáveis auxiliares qualitativas, a distribuição por classes na população. O IDEF ao inquirir ADP permite que se realize uma estratificação por características socioeconómicas e outras características de interesse. Os critérios de estratificação comuns são as regiões, fatores socioeconómicos do indivíduo de referência e tipo ou dimensão do agregado doméstico privado. A amostragem estratificada subdivide a população em classes segundo uma ou várias variáveis de estratificação. É possível através deste processo de amostragem a obtenção de estimativas com determinada precisão para certos grupos da população sem aumentos significativos de custos. Essas variáveis auxiliares, constituídas por informação para o total da população proveniente de fontes externas ao inquérito, são denominadas margens de calibração. Através da incorporação dessas margens nas fórmulas dos estimadores são modificados os ponderadores iniciais obtidos através do desenho da amostra. A calibração é um instrumento metodológico utilizado por institutos nacionais de estatística em grande parte dos inquéritos de forma a melhorar a qualidade dos estimadores. O método de calibração pretende corrigir os ponderadores iniciais, de modo a anular a sub-representação ou sobre representação, na amostra, de variáveis qualitativas e/ou quantitativas.

No IDEF, com o objetivo de assegurar a qualidade das observações para certas variáveis do rendimento e da despesa, foi necessário sobre dimensionar a amostra nas regiões de menor densidade populacional. Este facto originou desvios entre a estrutura regional da população revelada pelos Censos e a estrutura obtida a partir das respostas ao IDEF para certas características demográficas. De forma a serem corrigidos esses desvios, desde 1998, que o

INE utiliza o método de calibração de forma a corrigir os enviesamentos da amostra relativamente aos totais (conhecidos) da população. Assim no inquérito realizado em 2000 começaram-se a utilizar margens de calibração no cálculo dos ponderadores do IDEF através de informação auxiliar de diversas fontes de informação. A margem utilizada para o indivíduo foi a variável que representa o total de indivíduos por sexo cruzados com sete escalões etários (0-14 anos, 15-24 anos, 25-34 anos, 35-44 anos, 45-54 anos, 55-64 anos e mais de 65 anos). Em relação ao agregado, foi utilizada a variável correspondente ao total de ADP segunda a sua dimensão (1, 2, 3, 4, 5 e mais indivíduos). Foram também utilizadas margens para o alojamento, sendo utilizadas três variáveis: tipo de alojamento (moradia ou edifício de apartamentos), regime de ocupação (proprietário, arrendatário, subarrendatário, outro) e ano de construção (anterior a 1919, 1914-1945, 1946-1970, 1971-1990, após 1990). No inquérito realizado em 2005/06 as margens alteraram-se, mantendo-se as mesmas em 2010/11 e 2015/16. As margens utilizam de forma simultânea variáveis de agregado e de indivíduo: total de ADP por dimensão, total de ADP por região NUTS_II, total de ADP por nível de escolaridade do indivíduo de referência, total de indivíduos por escalão etário cruzado com o sexo e total de indivíduos por nível de escolaridade.

4.2. CONTAS NACIONAIS

A Contabilidade Nacional expõe de forma simplificada, coerente, agregada e quantificada as operações económicas efetuadas pelos agentes económicos num dado espaço geográfico (território económico) e num dado período de tempo. As Contas Nacionais (CN) são o produto resultante da Contabilidade Nacional.

As Contas Nacionais Portuguesas são produzidas e divulgadas pelo INE em conformidade com o Sistema Europeu de Contas Nacionais e Regionais (atualmente SEC 2010), que é consistente com o Sistema de Contas Nacionais das Nações Unidas (atualmente o SNA/System of National Accounts 2008). O SEC estabelece o quadro contabilístico da economia, das suas componentes e das suas relações com outras economias, constituindo um referencial metodológico de aplicação obrigatória na UE. Na elaboração das Contas

Nacionais é utilizada informação produzida pelo INE e por entidades externas (dados administrativos). As Contas Nacionais constituem um instrumento fundamental para a análise económica e para a formulação de políticas públicas.

O objetivo do trabalho prende-se com a criação de uma nova margem de calibração constituída através de informação auxiliar disponibilizada pelas Contas Nacionais. Desta forma, é realizada novamente a calibração para o IDEF 2015/16 com as margens já utilizadas e com a nova margem. A nova margem é uma variável que representa o total de despesa realizada por Divisão da COICOP (veja-se anexo 2 – figura 4).

É publicada, anualmente, informação do consumo privado (despesas de consumo final) das famílias no território económico por função COICOP ao segundo nível (Grupo) a preços correntes. Assim é possível utilizar a informação disponibilizada pelas Contas Nacionais como informação auxiliar para a calibração do IDEF.

4.3. REVISÃO BIBLIOGRÁFICA

No geral, a análise estatística quando aplicada a um conjunto de n unidades estatísticas e de p variáveis, y_1, y_2, \dots, y_p , considera como base uma matriz de dados $n \times p$, $[y_{ij}]$ $i = 1, 2, \dots, n$ $j = 1, 2, \dots, p$, onde cada linha, i , compreende as observações correspondentes a cada unidade inquirida (indivíduos, ADP, etc.) para todas as variáveis e cada coluna, j , contém as observações de cada uma das p variáveis para todas as unidades.

Depois do processo de recolha de dados as matrizes completas são uma exceção. A falta de resposta ou *missing data* é uma tendência que se tem vindo a agravar e afeta principalmente inquéritos sobre estudos sociais onde são inquiridos indivíduos, sendo um dos tipos de erro mais comum neste tipo de operações.

Biemer (2010) afirma que o erro total do inquérito é a “acumulação dos erros do desenho da amostra, recolha de dados, processamento e análise dos dados inquiridos”. Assim lacunas na matriz de dados podem dever-se a erros de amostragem, ou seja, *a priori*, ou a erros de não amostragem, ou seja, *a posteriori*. Os erros de não amostragem mais comuns são a não resposta, ou seja, a unidade selecionada depois de contactada recusa a sua participação no

inquérito, erros de cobertura ou erros de processamento de informação. Os erros de cobertura estão relacionados com a informação presente no FNA o que leva à existência de registos de alojamentos vagos, alojamentos não localizados e existência de novos alojamentos não registados.

A ocorrência de inconsistências em algumas variáveis leva também a uma necessidade de calibração das mesmas. Apesar dos inquéritos aos orçamentos familiares serem uma importante fonte de dados sobre a distribuição do consumo, existem bens para os quais o inquérito domiciliar fica muito significativamente aquém do total do consumo da população conhecido através de outras fontes estatísticas. Podemos conhecer essas diferenças através de informações auxiliares como análise de produção, importação, exportação, impostos especiais de consumo e outros inquéritos realizados. Deaton e Irish (1984) destacam esta problemática e indicam maiores diferenças entre os valores observados e os valores conhecidos para o total da população para bens alcoólicos e tabagistas. Os autores justificam esse acontecimento devido à exclusão da amostra de indivíduos que consomem elevadas quantidades desses bens e o facto dos indivíduos, quando questionados para esses consumos desviantes, omitirem os mesmos. A calibração, perante informação auxiliar para essas variáveis, permite a redução dessas disparidades.

Para Särndall (2007) a inexistência de erros de não amostragem é praticamente impossível, onde um dos principais problemas associados é o potencial enviesamento das estimativas do valor das variáveis para o total da população, quer sejam por calibração ou por outro método. Muitos artigos teóricos assumem a inexistência de não-resposta o que é na realidade uma situação muito rara. Como tal é importante examinar as discrepâncias geradas por esse erro e tentar reduzir o mesmo.

Thomsen e Siring (1980) retratam as consequências da não resposta através da análise de um inquérito acerca do comportamento eleitoral realizado pelas Estatísticas da Noruega (The Central Bureau of Statistics) realizado em 1969. A taxa de não resposta apresentada nesse estudo foi de 9.9%. No entanto, 81% dos indivíduos que recusaram responder votaram, assim como 65% dos indivíduos que não se encontravam no alojamento aquando do inquérito. Assim,

apesar de uma taxa de não resposta reduzida ocorreu uma sobre avaliação da taxa de resposta.

Existem grupos na população mais propensos a responder que outros. Um dos paradigmas é na comparação entre sexos, onde as mulheres, numa amostra aleatória simples, possuem uma maior representação nas respostas. As consequências são evidentes relativamente a diversos temas, por exemplo, se estivermos a questionar os rendimentos, irá existir um desvio em relação ao total, uma vez que as mulheres obtêm, na generalidade, um rendimento mais baixo (Särndal e Lundström 2005).

Hidirogrou et al. (1993) apresentam vários aspetos que conduzem a não resposta como o assunto do inquérito, métodos de recolha, período de realização, o acompanhamento do inquirido, entre outros. Gower (1979) destaca o papel do entrevistador e expõe que 15% dos entrevistadores não regista não respostas.

O IDEF realizado em 2015/2016 apresentou uma taxa de resposta de 64,1%. A elevada taxa de não resposta deve-se à carga estatística sobre os respondentes, uma vez que se trata de um inquérito longo e complexo, com muitas etapas de resposta.

O conceito de informação auxiliar é fundamental para a calibração. A informação auxiliar pode existir para cada um dos indivíduos da população (sexo e idade), para valores de parâmetros populacionais (média e totais) ou para a distribuição populacional de características que serão observadas na amostra mas não em todos os elementos da população. Existem situações onde a informação auxiliar tem uma estrutura complexa relacionada com o desenho da amostra. Em projetos onde o método de amostragem é multietápico a informação é composta por mais do que uma componente. Por exemplo, no caso de amostragem bietápica, a informação auxiliar pode existir apenas para as unidades primárias (Células GRID) ou unidades secundárias (alojamentos ou ADP). Consequentemente a calibração tem de ter em conta a estrutura da amostragem (Esteveo e Särndal, 2006). A existência de conjuntos de unidades para os quais não há dados amostrais mas para os quais temos informação

auxiliar, leva à necessidade do agrupamento de classes (Huang e Fuller, 1978; Park e Fuller, 2005 *cit. in* Estevao e Särndal, 2006).

Mohadjer et al., 1996, defendem o argumento de que a calibração é a forma conveniente para melhorar a eficiência das estimativas, utilizando informação externa. O autor apresenta várias vantagens na utilização da calibração:

- Equilíbrio: a informação fornecida pela amostra após calibração aproxima-se do valor total conhecido da população.
- Grupos-alvo: perante a necessidade de estimativas para uma subpopulação específica apenas é necessário adicionar a variável de calibração apropriada.
- Ancoragem de estimativas: uma variável correlacionada com uma variável de calibração não se distanciará, para valores longínquos, fornecendo estabilidade ao longo do tempo.
- Consistência: após a calibração cada unidade da amostra admite apenas um ponderador final.

A calibração promove a coerência, uma vez que se obtém a mesma estimativa, independentemente do inquérito, para a mesma variável.

Os ponderadores dos inquéritos são um elemento importante na estimativa das características da população, porque permitem obter estimativas mais próximas dos respetivos valores populacionais. Os ponderadores calibrados para as estimativas dos totais coincidem com os totais populacionais. O processo de ponderação começa na seleção da amostra, dado que cada unidade da população é selecionada para fazer parte da amostra com determinada probabilidade (Shaffer et al., 2014). Se todas as unidades da amostra tiverem a mesma probabilidade de seleção todas as unidades terão o mesmo ponderador no projeto.

Horwitz & Thompson (1952) criaram o estimador (de agora em diante designado por estimador HT), de forma a estimar o total de Y para toda a população, dado por

$$(1) \quad \hat{\theta} = \sum_{k \in S} w_k Y_k$$

onde w_k , ponderador amostral de cada elemento da amostra, é o inverso da probabilidade de seleção, π_k , ($w_k = 1/\pi_k$) do mesmo, S o conjunto das unidades da amostra, e Y_k o valor da variável Y para a unidade k da amostra.

O estimador de HT conduziu à aceitação das técnicas de ponderação. Hoje é utilizado como estimador inicial na calibração de muitos inquéritos. O IDEF utiliza também um ponderador baseado no estimador de HT como ponderador inicial, contudo adaptado para contemplar a problemática da não resposta. Na presença de não resposta o estimador é enviesado, sendo então, necessário a utilização de um estimador revisto.

A probabilidade da k -ésima unidade pertencer à amostra e responder é igual a $\pi_k \varphi_k$ onde φ_k é a propensão à resposta. A probabilidade φ_k tem de ser estimada para cada uma das unidades finais, com base em variáveis conhecidas para a população. Assim, o peso calibrado para cada uma das unidades que responderam é $\tilde{w}_k = \frac{w_k}{\hat{\varphi}_k} = \frac{1}{\hat{\varphi}_k * \pi_k}$. Os estimadores são calculados com base nas respostas

$$(2) \quad \hat{\theta} = \sum_{k \in Resp} \tilde{w}_k Y_k$$

onde $Resp$ é o conjunto das unidades da amostra que responderam ao inquérito.

A precisão associada a um estimador $\hat{\theta}$ para o parâmetro θ (geralmente desconhecido), pode ser medida em termos absolutos pelo inverso da variância ou desvio padrão e em termos relativos pelo coeficiente de variação igual a,

$$(3) \quad CV(\hat{\theta}) = \frac{\sqrt{\hat{var}(\hat{\theta})}}{\hat{\theta}} * 100\%$$

Por definição, os ponderadores iniciais (inverso da probabilidade de seleção) são maiores ou iguais a 1, ou seja, cada unidade observada representa um conjunto de unidades na população. Porém, os ponderadores calibrados não são necessariamente maiores ou iguais a 1, a menos que seja tomado cuidado especial no seu cálculo. É importante na construção de ponderadores através da calibração estabelecer limites aos ponderadores de forma a evitar ponderadores negativos e ponderadores demasiado elevados. Isto pode ser alcançado através

de uma escolha adequada da função distância, entre o ponderador final e o ponderador inicial, subjacente à calibração.

5. CALIBRAÇÃO ATRAVÉS DA MACRO CALJACK

O software utilizado pelo INE no IDEF é o *SAS – Statistical Analysis Software*. O SAS é um sistema integrado de aplicações para a análise de dados que permite a recuperação de dados, a gestão de ficheiros, análises estatísticas e a obtenção de gráficos e relatórios. O SAS nasceu de um consórcio de oito universidades, nos EUA devido a uma bolsa criada pelo *National Institutes of Health (NIH)* com o propósito de analisar uma enorme quantidade de dados agrícolas. A contínua procura e descoberta de áreas em que seria possível a aplicação do SAS levou à sua presença, atualmente, em mais de 146 países.

O INE utiliza, para a realização da calibração, a macro CALJACK desenvolvida em SAS. A macro foi cedida por Pierre Lavallée das Estatísticas do Canadá e é uma modificação da macro CALMAR elaborada por Sautory do INSEE (França). A macro CALJACK com as margens criadas com informação auxiliar, extrapola para o total da população, os valores observados na amostra, para variáveis qualitativas e/ou quantitativas. No CALJACK é possível o cálculo de um estimador para a população do total das variáveis observadas da amostra, do quociente de totais ou da diferença de totais entre duas variáveis ou da diferença de quocientes, utilizando um estimador da variância do tipo Jackknife. No IDEF é realizada a calibração para os totais das variáveis anteriormente referidas (margens) e para o rácio da despesa, realizada por COICOP ou agrupamento de classes da COICOP, por ADP.

Como referido anteriormente, a calibração consiste na substituição dos ponderadores iniciais, obtidos através da definição do método de amostragem, pelos ponderadores finais calibrados, tão próximos quanto possível dos iniciais, utilizando uma função distância. A função distância tem de garantir que o total das variáveis (qualitativas e/ou quantitativas), extrapoladas com os ponderadores calibrados, seja igual ao total populacional conhecido. A solução teórica do problema passa pela escolha de uma função distância, que sujeita às condições de calibração, minimize as distâncias entre os ponderadores iniciais

e os ponderadores finais calibrados de forma a controlar a variância das estimativas, e assim, garantir um certo nível de qualidade das mesmas.

Considere-se que o IDEF tem por base uma amostra de dimensão s selecionada a partir de uma população de dimensão N , $U = \{1 \dots k \dots N\}$. Para todo o alojamento k de U , $\pi_{k|jh}$ é a probabilidade da sua inclusão em s para cada unidade primária j do estrato h (NUTS_II).

A seleção do alojamento é realizada em duas fases: seleciona-se na primeira etapa a PSU j e na segunda etapa dentro da PSU j seleciona-se o alojamento k . A probabilidade de seleção de cada alojamento k ($\pi_{k|jh}$) é calculada da seguinte forma

$$(4) \quad \pi_{k|jh} = \begin{cases} \frac{A_{jh}}{I_h} * \frac{n_{jh}}{A_{jh}} = n_{jh} * \frac{n_h}{A_h} & \text{se } A_{jh} < I_h \\ \frac{n_{jh}}{A_{jh}} & \text{se C.C.} \end{cases}$$

com

$$(5) \quad I_h = \frac{A_h}{n_h}$$

$\pi_{k|jh}$ - Probabilidade (condicionada) de seleção do alojamento k pertencente à PSU j no estrato h ;

A_{jh} - Número de alojamentos de residência principal da PSU j no estrato h ;

I_h - Passo de seleção sistemática da 1ª etapa no estrato h ;

n_h - Número de PSU selecionadas no estrato h ;

n_{jh} - Número de alojamentos de residência principal selecionados na PSU j no estrato h ;

A_h - Número de alojamentos de residência principal no estrato h .

Seja y_k o valor, para o alojamento k , de uma variável Y . O total de Y estimado para a população, utilizando o estimador de HT é

$$(6) \quad \hat{\theta} = \sum_{k \in U} \frac{1}{\hat{\varphi}_k * \pi_k} y_k = \sum_{k \in S} d_k y_k$$

onde $\hat{\varphi}_k$ é propensão à resposta, que é estimada para cada uma das unidades finais, com base em variáveis conhecidas para a população e

$$(7) \quad d_k = \frac{1}{\hat{\varphi}_k * \pi_k}$$

A utilização deste estimador permite afetar, a cada indivíduo da amostra, um ponderador d_k igual ao inverso da probabilidade de seleção tendo em conta a propensão à resposta.

Sejam $x_1 \dots x_j \dots x_J$, J variáveis auxiliares observadas na s , para as quais se conhecem os totais da população X_j .

Para ter em conta esta informação, estima-se θ , o total de Y por intermédio de um estimador da forma

$$(8) \quad \widehat{\theta}_w = \sum_{k \in s} w_k y_k$$

em que os ponderadores finais calibrados w_k , afetos aos indivíduos são tão próximos quanto possível dos ponderadores iniciais d_k , no sentido de uma função específica de distância a definir, que verifica as equações de calibração

$$(9) \quad \sum_{k \in s} w_k x_{jk} = X_j, \forall j = 1 \dots J$$

Seja G a função de distância escolhida, os ponderadores w_k ($k \in s$) são soluções de

$$(10) \quad \text{Min} \sum_{k \in s} d_k G\left(\frac{w_k}{d_k}\right)$$

sob as restrições de ajustamento

$$(11) \quad \sum_{k \in s} w_k x_k = X$$

Ou seja, minimiza-se uma soma ponderada (pelos d_k) das distâncias entre os ponderadores de amostragem d_k e as ponderações procuradas w_k , sob as restrições de calibração.

O método utilizado pelo INE na calibração da amostra do IDEF para a função distância é o método logit. O método logit adquire esta designação porque a função inversa da derivada da função $G(\cdot)$ é a função logística. Para realizar a minimização é utilizado o método iterativo de Newton. A função $G(\cdot)$ depende do parâmetro L e do parâmetro U . Não se pode escolher *a priori* o valor para L e U fazendo-se por aproximações sucessivas, diminuindo progressivamente L e aumentando progressivamente o U . No entanto, é definido um valor máximo

(L_{\max} inferior a 1) e um valor mínimo (U_{\min} superior a 1). O INE definiu também um valor máximo para U igual a 30 e um valor mínimo para L igual a 0.03 ($\approx 1/30$).

A macro CALJACK permite realizar essa extrapolação da amostra e estimar a variância dos valores ajustados das variáveis em estudo. Os ponderadores finais ajustados quando aplicados a cada variável auxiliar produzem estimativas cujo total coincide com o total populacional de cada categoria da população considerada.

O método Jackknife é aconselhável para calcular a variância de estimadores não lineares, como é o caso de estimadores baseados numa ponderação de ajustamento por margens em que não existe nenhuma fórmula específica de variância. Para o cálculo da variância, a partir da amostra são geradas g amostras alternativas, denominadas de réplicas. O número de réplicas deve ser escolhido ponderando o tempo de cálculo (aumenta com g) e a precisão (aumenta com g). O estimador da variância é dado por

$$(12) \quad \widehat{var}(\hat{\theta}) = \frac{(g-1)}{g} \sum_{m=1}^g (\hat{\theta}_m - \hat{\theta})^2$$

onde $\hat{\theta}$ é um estimador de θ (calculado sobre a amostra total) e $\hat{\theta}_m$ o estimador de θ para a réplica m .

O método de calibração por margens é realizado para cada região geográfica (NUTS_II) utilizando-se margens para o indivíduo e para o agregado. Recorde-se que serão utilizadas as margens já empregues pelo IDEF e a nova margem construída com a informação disponibilizada pelas CN.

O primeiro passo para a realização da calibração utilizando a variável auxiliar disponibilizada pelas CN é a harmonização da informação entre as CN e o IDEF para que seja possível a comparação das mesmas. As CN apresentam apenas os consumos realizados pelos ADP em território nacional por classe da COICOP. O IDEF apresenta informação das despesas de consumo, realizadas em Portugal e no estrangeiro, na sua totalidade. Porém, é recolhida informação que permite a criação por parte do INE dessa distinção, uma vez que no inquérito é questionado, para cada consumo, se o mesmo foi realizado em território nacional ou não. Assim é possível a construção de variáveis de consumo equiparáveis às das CN.

O ficheiro “estrangeiro” possui informação acerca do consumo realizado fora do território nacional, por COICOP a primeiro nível. No ficheiro “idef.idef15”, disponível pelo Departamento de Metodologia e Sistemas de Informação (DMSI) está disponível informação para todas as variáveis necessárias à calibração. Em ambos os ficheiros, os ADP são identificados através do código do agregado dado pela variável *Idagregado*, o que permite a reunião dos dois ficheiros em um outro intitulado “novo”. Após a criação desse ficheiro é possível criar novas variáveis que resultam da diferença entre o consumo total e o consumo realizado no estrangeiro para cada COICOP a primeiro nível. Os comandos utilizados para a criação do ficheiro novo podem ser consultados em anexo (veja-se anexo 2 – figura 1).

Para observar o efeito da nova calibração foi calculado a correlação, entre a série com o total das variáveis calibradas por COICOP do IDEF e a série com o total das mesmas variáveis por COICOP das CN. A forte correlação ($=0.8779$) deverá garantir um bom comportamento dos ponderadores relativamente às variáveis em estudo.

Segundo o IDEF e as CN os ADP praticam as suas maiores despesas nas Divisões 01 (produtos alimentares e bebidas não alcoólicas), 04 (habitação, água, eletricidade, gás e outros combustíveis) e 07 (Transportes) da COICOP (veja-se anexo 1). A tabela 1 apresenta as diferenças para as Divisões 01, 04 e 07 obtidas através das observações do IDEF calibradas com o método realizado pelo INE em 2015/16 e das CN.

Tabela 1 – Comparação entre os totais das variáveis COICOP para diferentes fontes de informação.

Nível da COICOP	IDEF (10⁶)	CN (10⁶)
04	26 369	23 116
01	11 942	20 817
07	11 414	14 996

Nota: IDEF – observações, retiradas do IDEF, calibradas; CN – Contas Nacionais

O INE calcula os ponderadores iniciais para cada ADP para cada NUTS_II. Assim no ficheiro “idef15” reunimos os ponderadores iniciais obtidos através do método de amostragem, presentes no ficheiro “ponderadores”, de todos os ADP e a informação reunida no ficheiro “novo”. A união destes ficheiros é possível através do cruzamento da informação da variável *Idagregado* presente no

ficheiro “novo” com a variável *idadp* presente no ficheiro dos “ponderadores” (veja-se anexo 2 – figura 2).

De seguida, através da variável *NUTS_2* que contém os códigos das regiões geográficas são criados 7 ficheiros, um para cada uma das *NUTS_II*. Cada ficheiro possui apenas os ADP dessa *NUTS_II*. Os comandos utilizados para a criação desses ficheiros podem ser consultados em anexo (veja-se anexo 2 – figura 2). A macro *CALJACK* é aplicada individualmente a cada região da *NUTS_II*. Assim para cada uma delas é necessário definir todos os parâmetros necessários à calibração (veja-se anexo 3).

No processo de calibração, é definido para cada região diferentes valores de *LO* e *UP* e são criadas diferentes variáveis devido à necessidade de agrupamento de classes *COICOP*. A variável *desj* define a despesa realizada na divisão da *COICOP* *j* pelos ADP da *NUTS_II* com $j = 1, \dots, 12$.

Para a região Norte realiza-se a calibração de todas as variáveis quantitativas e qualitativas através do valor 0,07 ($\approx 1/15$) para o parâmetro *LO* e com o valor de 15 para o parâmetro *UP* (veja-se anexo 4 figura 5).

Na região Centro não é possível a calibração para todas as margens. Procede-se assim ao agrupamento de divisões da *COICOP*. Através do SAS é possível criar novas variáveis. A variável “des16” resulta da soma entre as variáveis “des09” e “des06”, a variável “des17” resulta do agrupamento das variáveis “des05” e “des02” e por fim a variável “des18” é a junção das variáveis “des08” e “des10”. Assim, foi possível estabelecer para o parâmetro *UP* o valor de 25 e para o parâmetro *LO* o valor de 0,04 ($= 1/25$) e realizar a calibração (veja-se anexo 4 – figuras 6 e 7).

A calibração, na região da Área Metropolitana de Lisboa, foi possível para todas as variáveis sem necessidade de agrupamento, com um valor de 0,045 para *LO* e 22 para *UP*. (veja-se anexo 4 – figura 8).

Para o Alentejo a calibração de todas as variáveis quantitativas e qualitativas não foi possível sem se realizar o agrupamento de divisões. Só conseguimos calibrar de forma independente para as variáveis “des04” e “des01”, assim foi criada uma nova variável “des26”, que resulta da soma das restantes variáveis.

O valor estabelecido para o parâmetro LO foi 0,03 ($\approx 1/30$) para o parâmetro UP foi 30 (veja-se anexo 4 – figuras 9 e 10).

A calibração através da informação auxiliar das Contas Nacionais foi realizada com sucesso para todas as variáveis para o Algarve. Os valores definidos para UP e LO foram, correspondentemente, 25 e 0,04 ($=1/25$) (veja-se anexo 4 – figura 11).

Para a Região Autónoma dos Açores foi também realizado um agrupamento de divisões. A calibração deu-se então para as variáveis “des01”, “des04” e “des27”, sendo a variável “des27” igual à soma das restantes variáveis. O valor final do parâmetro LO foi 0,03 e do UP foi 30 (veja-se anexo 4 – figuras 12 e 13).

Para a Região Autónoma da Madeira foram agrupadas todas as variáveis duas a duas por ordem decrescente do seu valor. Assim foram definidas seis novas variáveis: “des13”, “des14”, “des23”, “des24”, “des25”, “des26”. Os valores de UP e LO foram 15 e 0,07, correspondentemente (veja-se anexo 4 – figuras 14 e 15).

6. APRESENTAÇÃO E ANÁLISE DE RESULTADOS

Ao utilizar como informação auxiliar as Contas Nacionais, pretende-se uma aproximação das estimativas do IDEF aos totais que constam nas CN. A estrutura da despesa total anual média por agregado em 2015, por divisão da COICOP e NUTS_II, permite identificar a afetação da despesa dos ADP nesse ano. Uma vez que as CN apenas fornecem informação acerca das despesas efetuadas pelo ADP em território nacional, ao longo do trabalho foi sempre usada informação apresentada pelo IDEF sem as despesas realizadas no estrangeiro. Assim é importante a comparação entre a estrutura da despesa do ADP segundo as CN (tabela 2), o IDEF com a calibração original (tabela 3) e o IDEF com a nova calibração que incorpora as CN como informação auxiliar (tabela 4).

Tabela 2 – Estrutura da despesa total anual média por ADP em 2015, por divisão da COICOP segundo as Contas Nacionais, em percentagem.

Classes da COICOP	Território Nacional
01	17,1
02	3,1
03	6,4

04	19,0
05	5,1
06	5,1
07	12,3
08	2,4
09	6,1
10	1,2
11	11,8
12	10,4

Fonte: site oficial do INE – Contas Nacionais

Nota: Classes da COICOP a primeiro nível (veja-se anexo 1).

A estrutura do consumo apresentada pelas CN (tabela 2) difere bastante da estrutura apresentada pelo IDEF (tabela 3). Podemos observar as divisões da COICOP de maior consumo. A divisão 01 (produtos alimentares e bebidas não alcoólicas) apresenta um peso de 17% nas despesas do ADP segundo as CN enquanto segundo o IDEF apresentava um peso de 14,6%. A divisão 04 (habitação e manutenção da mesma) é a divisão que apresentava maiores diferenças, sendo o peso de 19% para as CN e superior a 30% em todas as regiões NUTS_II de acordo com o IDEF.

Tabela 3 – Estrutura da despesa total anual média por ADP, realizada em território nacional, por divisões da COICOP, NUTS_II, em 2015/16 (calibração sem informação auxiliar das CN) em percentagem.

COICOP	Portugal	Norte	Centro	A.M. Lisboa	Alentejo	Algarve	R.A. Açores	R.A. Madeira
01	14,6	14,7	14,6	13,5	18,0	15,4	15,4	14
02	1,6	1,8	1,3	1,2	2,0	2,1	2,3	2,0
03	3,4	3,9	3,1	3,1	3,0	3,4	2,6	3,0
04	32,2	30,5	33,7	33,1	33,0	31,7	37,9	34,0
05	4,0	4,0	4,3	4,1	3,0	3,2	4,5	4,0
06	5,6	5,4	6,3	5,4	6,0	5,4	6,8	6,0
07	13,9	14,0	15,0	13,9	13,0	12,9	12,1	14,0
08	3,2	3,2	3,2	3,2	4,0	3,2	4,3	4,0
09	3,9	4,0	3,3	4,6	3,0	3,8	3,5	3,0
10	2,1	2,2	1,6	2,8	1,0	1,4	1,2	2,0
11	8,6	9,1	7,1	8,2	9,0	10,8	4,6	8,0
12	6,9	7,2	6,9	6,9	6,0	6,9	5,2	6,0

Fonte: site oficial do INE

Nota: Classes da COICOP a primeiro nível (veja-se anexo 1).

Com a realização da nova calibração verificam-se alterações na estrutura de despesa para todas as divisões da COICOP, aproximando-se assim os valores à estrutura apresentada pelas CN, como seria de esperar e se pode observar na tabela 4. A classe 04 passou a ter um peso, no geral de 19% para todas as regiões da NUTS_II, exceto a R.A. da Madeira que apresenta um valor de 24,9% e a classe 01 passou a apresentar valores próximo de 17% com exceção da R.A. da Madeira (11%).

Tabela 4 – Estrutura da despesa total anual média por agregado, realizada no território nacional, por classe da COICOP, NUTS_II, em 2015/16 (calibração com informação auxiliar das CN), em percentagem.

COICOP	Portugal	Norte	Centro	A.M. Lisboa	Alentejo	Algarve	R.A. Açores	R.A. Madeira
01	16,2	17,1	17,0	16,9	17	17	17,1	11,0
03	2,2	3,1	1,3	3,1	1,8	3,1	1,6	1,3
02	5,5	6,5	6,4	6,6	3,1	6,5	3,3	6,2
04	19,8	19,0	18,9	18,8	19,0	18,9	19,0	24,9
05	5,5	5,1	6,9	5,1	3,7	5,1	5,8	6,9
06	5,8	5,1	6,5	5,1	5,4	5,2	6,5	6,5
07	16,9	12,3	12,3	12,3	28,7	12,3	25,1	15,3
08	2,6	2,6	2,6	2,4	2,4	2,4	3,1	2,8
09	5,0	6,1	4,7	6,1	3,0	6,1	4,2	4,6
10	1,2	1,2	1,0	1,2	1,5	1,2	1,6	0,7
11	10,4	11,8	12	12,3	9,1	11,7	7,2	8,7
12	9,1	10,5	10,5	10,4	5,3	10,7	5,4	11,1

Nota: Classes da COICOP a primeiro nível (veja-se anexo 1).

De forma a analisar a qualidade dos estimadores e a precisão dos mesmos é necessário observar o respetivo coeficiente de variação. O CV é utilizado para analisar a dispersão dos estimadores face aos seus valores médios. Assim, quanto menor o CV maior será a precisão dos estimadores.

Os coeficientes de variação dos estimadores para as variáveis das margens já utilizadas não variaram muito com a nova calibração, mantendo-se em valores aproximadamente iguais a zero.

É essencial a comparação dos CV dos estimadores dos rácios, ou seja, dos rácios da despesa média por ADP por nível da COICOP. Ao correr a macro

CALJACK é criado para cada região geográfica da NUTS_II um ficheiro com esses coeficientes de variação. Os coeficientes de variação para esses estimadores melhoraram bastante depois da introdução da informação auxiliar na calibração. Quando a calibração era realizada apenas para as cinco margens existentes o coeficiente de variação do rácio apresentava valores compreendidos entre 1 e 13. Após a calibração com a nova margem construída com a informação auxiliar das CN o CV passou a exibir valores próximos de 0, salvo muito raras exceções (veja-se anexo 5). Esta diminuição do coeficiente de variação representa um aumento da precisão das estimativas.

Contudo, apesar desta melhoria da qualidade é necessário refletir acerca de outros aspetos importantes.

Para que a calibração fosse possível teve de se proceder ao agrupamento de divisões da COICOP na maior parte das regiões NUTS_II. O principal aspeto que influencia a necessidade do agrupamento de classes é a pouca densidade populacional da respetiva NUTS_II, existindo assim maiores agrupamentos de classes nas regiões autónomas.

Ao longo da pesquisa acerca da informação prestada pelas Contas Nacionais e a sua comparação com o IDEF, foi perceptível uma enorme diferença de valores na divisão 4 da COICOP que corresponde aos custos com a habitação, água, eletricidade, gás e outros combustíveis. Segundo o IDEF 2015/16, antes de realizada a nova calibração, essa classe apresentava um peso no consumo dos ADP entre 30,7%, para a região Norte, e os 37,9%, para a Região Autónoma dos Açores, enquanto para as CN apresentava para o total do território um peso de 19%. Esta diferença apresentada entre as duas fontes de informação deve-se à subclasse 04.2 da COICOP relacionada com o arrendamento fictício da habitação. No IDEF é utilizada a auto-locação para situações de agregados que são proprietários da residência, arrendatários a preço abaixo do mercado ou em situações de usufruto gratuito ou a título de salário. A auto-locação corresponde a uma estimativa subjetiva obtida pelo próprio ADP sobre o valor hipotético de renda do seu alojamento a preços de mercado. Para tal, o ADP deve ter em conta a localização da habitação, a área do alojamento, o número de divisões, o estado de conservação entre outros aspetos. Perante as mesmas situações, as

CN imputam o valor, sendo utilizado o método de estratificação. O parque habitacional é estratificado tendo em conta a localização, natureza da habitação e outros fatores influenciadores da renda. Para obter uma estimativa do valor locativo do total do parque habitacional são utilizadas informações relativas a rendas efetivas de habitações arrendadas. A renda efetiva média por estrato é aplicada a todas as habitações desse estrato particular. Esses valores podem ser corrigidos tendo em conta informação relativamente ao mercado imobiliário e a análises de regressão. Estas diferenças explicam assim o diferente peso dessa área de consumo entre o IDEF e as CN e conseqüentemente das restantes classes da COICOP, uma vez que as despesas com a habitação têm um grande impacto nas despesas dos ADP.

Outro aspeto a ter em conta na análise dos resultados obtidos é a origem das fontes da informação disponibilizada pelas CN. As CN utilizam informação produzida pelo INE, por entidades externas produtoras de estatísticas oficiais (por delegação de competências do INE) e por outras entidades externas através de dados administrativos. As CN utilizam assim informação recolhida por vários inquéritos realizados no INE sendo o IDEF um dos inquéritos utilizado como fonte de informação. As CN ao utilizarem essa informação no cálculo do consumo privado (despesa de consumo final) das famílias no território económico por função (COICOP), pode tornar incorreta a sua utilização como informação auxiliar na calibração dos estimadores do IDEF.

7. CONCLUSÃO

Este trabalho final de mestrado teve como objetivo testar métodos alternativos de calibração do IDEF. Para tal usou-se informação adicional auxiliar proveniente das CN.

O objetivo do trabalho surge com a necessidade de corrigir os efeitos resultantes da amostragem, das não respostas e de outros erros que afetam os inquéritos. A calibração é o método mais eficiente e coerente para a correção desses mesmos erros, *a posteriori*.

De forma a desenvolver este projeto foi utilizado o *software* SAS através da aplicação da macro CALJACK. A macro CALJACK permitiu assim realizar a calibração para diversos inquéritos sendo utilizada por vários Institutos Nacionais

de Estatística. Foi inserida a nova informação auxiliar, ou seja, uma nova margem de calibração, de forma a aproximar os valores registados no inquérito da realidade conhecida para a população.

O novo método de calibração permitiu uma melhoria da qualidade das observações do IDEF calibradas, avaliada através dos coeficientes de variação. Depois de realizada a calibração, tal como era objetivo, a estrutura de consumo dos ADP aproximou-se da estrutura de consumo apresentada pelas CN.

Para pesquisas futuras é sugerido a realização da imputação ao invés da auto-locação para a subclasse 04.2 da COICOP (rendas subjetivas pela habitação) de forma a aproximar os valores entre as CN e o IDEF antes de realizada a calibração. De forma a utilizar novamente a informação das CN para a realização da calibração do IDEF é necessário ter em conta o calendário do INE, uma vez que a informação das CN tem de estar disponível aquando da realização da calibração para o IDEF.

No geral, o objetivo do trabalho foi concretizado, as expectativas foram correspondidas e a realização do mesmo foi um processo estimulante.

REFERÊNCIAS

- Biemer, P.P. (2010). Total Survey Error: Design, Implementation, and Evaluation. *American Association for Public Opinion Quarterly*, Vol.74, No. 5, pp.817-848. Oxford University Press. Reino Unido.
- Coelho, P., Pereira, L., Pinheiro, J. e Xufre, P. (2017). *As Sondagens: Princípios, Metodologias e Aplicações*. 1ª Edição, Escolar Editora. Lisboa.
- Costa Lorga da Silva, A.I.G. (1995). *Falta de Resposta em Sondagens*. Dissertação de mestrado em Matemática Aplicada à Economia e à Gestão. Instituto Superior de Economia e Gestão – Universidade de Lisboa. Lisboa.
- Deaton, A. and Irish, M. (1984). Statistical Models for Zero Expenditures in Household Budgets. *Journal of Public Economics*, pp.59-80. North-Holland.
- Departamento de Estatísticas Demográficas e Sociais (2015). *Documento Metodológico*. Versão 2.0., Instituto Nacional de Estatística. Lisboa.
- Departamento de Estatísticas Demográficas e Sociais (2017). *Inquérito às Despesas das Famílias 2015/2016*. Edição 2017, Instituto Nacional de Estatística. Lisboa.
- Departamento de Metodologia e Sistemas de Informação (2015). *Sistema Global de Gestão de Inquéritos por Entrevista – SIGINQ – IE*. Enquadramento para o FNA. Instituto Nacional de Estatística. Lisboa.
- Estevao, V.M. and Särndal, C.-E. (2006). Survey Estimates by Calibration on Complex Auxiliary Information. *International Statistical Review*, Vol. 74, pp. 127-147. Canada.
- Gower, A.R. (1979). Nonresponse in the Canadian Labour Force Survey. *Survey Methodology – Statistics Canada*. Canada.
- Hidiroglou, M., Douglas, J. and Gray, G. (1993). *A Framework for Measuring and Reducing Nonresponse in Surveys*, Vol. 19, No.1, pp.81-94. Statistics Canada. Canada.

- Horwitz, D.G., and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*. Vol. 47, No. 260, pp. 663- 685.
- Huang, E.T. and Fuller, W.A. (1978). Nonnegative regressions estimation for sample survey data. *Proceedings Social Statistics Section – American Statistical Association*.
- Instituto Nacional de Estatística (2014) Carta da Qualidade. 4ªEdição, Instituto Nacional de Estatística. Lisboa.
- Lohr, S.L. (2010). *Sampling: Design and Analysis*. 2nd Edition, Brooks/Cole. USA.
- Mohadjer, L., Yansaneh, I. and Brick, M. (1996). Application of logistic models to survey data using replication techniques. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, pp. 462–467. Washington, D.C.
- Office for Official Publications of the European Communities. (2003) *Household Budget Surveys in the EU*. Methodology and recommendations for harmonization. European Communities. Luxembourg.
- Park, M. & Fuller, W.A. (2005). Towards nonnegative regression weights for survey samples. *Survey Methodology*, Catalogue No. 12-001, Vol.31, No.1, pp.85-93. Statistics Canada.
- Särndal, C.E. (2007). The calibration approach in survey theory and practice, Vol 33, No. 2, pp. 99-119. Catalogue no. 12-001-X, *Statistics Canada*. Canada.
- Särndal, C.E. and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. 1nd Edition, John Wiley & Sons. EUA.
- Shaffer, B., Cheng, Y. and Slud, E. (2014). *Single-Stage Generalized Raking Application in the American Housing Survey*. U.S. pp. 3708-3719. Census Bureau.
- Thomsen, I. and Siring, Erling (1980). *On the Causes and Effects on Non-response: Norwegian Experiences*. Statistics Norway. Oslo.

ANEXOS

**ANEXO 1 – COICOP: CLASSIFICAÇÃO INDIVIDUAL DO CONSUMO DOS ADP
POR FUNÇÃO**

Tabela 5 – Níveis da COICOP

COICOP - Classificação individual do consumo dos ADP por função			
Código COICOP			Descrição
1ºNível	2ºNível	3ºNível	
01			Produtos alimentares e bebidas não alcoólicas
	01.1		Produtos alimentares
		01.1.1	Pão e cereais
		01.1.2	Carne
		01.1.3	Peixe, crustáceos e moluscos
		01.1.4	Leite, queijo e ovos
		01.1.5	Óleos e gorduras
		01.1.6	Fruta
		01.1.7	Produtos hortícolas
		01.1.8	Açúcar, confeitaria, mel e outros produtos à base de açúcar
		01.1.9	Produtos alimentar i.e.
	01.2		Bebidas não alcoólicas
		01.2.1	Café, chá e cacau
		01.2.2	Águas minerais, refrigerantes e sumos de fruta e de produtos hortícolas
02			Bebidas alcoólicas, tabaco e narcóticos
	02.1		Bebidas alcoólicas
		02.1.1	Bebidas espirituosas
		02.1.2	Vinho
		02.1.3	Cerveja
	02.2		Tabaco
	02.3		Narcóticos
03			Vestuário e calçado
	03.1		Vestuário
		03.1.1	Tecidos para vestuário
		03.1.2	Artigos de vestuário
		03.1.3	Outros artigos e acessórios de vestuário
		03.1.4	Limpeza, reparação e aluguer de vestuário
	03.2		Calçado
		03.2.1	Calçado
		03.2.2	Reparação e aluguer de calçado
04			Habitação, água, eletricidade, gás e outros combustíveis
	04.1		Rendas efetivas pela habitação
		04.1.1	Rendas efetivas de residência principal pagas por inquilinos
		04.1.2	Outras rendas efetivas pagas por inquilinos
	04.2		Rendas subjetivas (arrendamento fictício) pela habitação
		04.2.1	Rendas subjetivas de residência principal (proprietários)
		04.2.2	Outras rendas subjetivas
	04.3		Manutenção e reparação das habitações
		04.3.1	Materiais para a manutenção e reparação da habitação
		04.3.2	Serviços para a manutenção e reparação da habitação
	04.4		Abastecimento de água e outros serviços relacionados com a habitação
		04.4.1	Abastecimento de água
		04.4.2	Recolha de resíduos sólidos

COICOP - Classificação individual do consumo dos ADP por função			
Código COICOP			Descrição
1ºNível	2ºNível	3ºNível	
		04.4.3	Saneamento básico
		04.4.4	Outros serviços relacionados com a habitação, n.d.
	04.5		Eletricidade, gás e outros combustíveis
		04.5.1	Eletricidade
		04.5.2	Gás
		04.5.3	Combustíveis líquidos
		04.5.4	Combustíveis sólidos (para aquecimento e iluminação)
		04.5.5	Energia térmica
05			Mobiliário, artigos de decoração, equipamento doméstico e manutenção corrente da habitação
	05.1		Mobiliário, artigos de iluminação, decoração, tapetes e outros revestimentos para o chão
		05.1.1	Mobiliário e acessórios para o lar
		05.1.2	Carpetes e outros revestimentos para pavimentos
		05.1.3	Reparação de mobiliário, acessórios para o lar, carpetes e outros revestimentos para pavimentos
	05.2		Têxteis de uso doméstico
	05.3		Equipamento doméstico
		05.3.1	Grandes aparelhos eletrodomésticos
		05.3.2	Pequenos eletrodomésticos
		05.3.3	Reparação de equipamento doméstico
	05.4		Vidros, louças e outros utensílios de uso doméstico
	05.5		Ferramentas e equipamento para casa e jardim
		05.5.1	Grandes ferramentas e equipamentos para casa e jardim
		05.5.2	Pequenas ferramentas e acessórios diversos
	05.6		Bens e serviços para a manutenção corrente da habitação
		05.6.1	Bens não duráveis de uso doméstico
		05.6.2	Serviços domésticos e serviços relativos à habitação
06			Saúde
	06.1		Produtos, aparelhos e equipamentos médicos
		06.1.1	Medicamentos e especialidades farmacêuticas
		06.1.2	Outros produtos médicos e farmacêuticos
		06.1.3	Aparelhos e material terapêutico
	06.2		Serviços médicos, paramédicos e outros serviços de saúde não hospitalares
		06.2.1	Serviços médicos
		06.2.2	Serviços de medicina dentária
		06.2.3	Serviços Paramédicos
	06.3		Serviços hospitalares
07			Transportes
	07.1		Aquisição de veículos
		07.1.1	Aquisição de veículos automóveis
		07.1.2	Aquisição de ciclomotores e motociclos
		07.1.3	Aquisição de bicicletas (sem motor)
		07.1.4	Aquisição de veículos de tração animal
	07.2		Despesas com a utilização de veículos para transporte pessoal
		07.2.1	Peças e acessórios para equipamento de transporte pessoal
		07.2.2	Combustíveis e lubrificantes para equipamento de transporte pessoal
		07.2.3	Manutenção e reparação de equipamento de transporte pessoal

COICOP - Classificação individual do consumo dos ADP por função			
Código COICOP			Descrição
1ºNível	2ºNível	3ºNível	
		07.2.4	Outros serviços relacionados com o equipamento de transporte pessoal
	07.3		Serviços de transportes
		07.3.1	Transportes ferroviários de passageiros
		07.3.2	Transportes rodoviários de passageiros
		07.3.3	Transportes aéreos de passageiros
		07.3.4	Transportes de passageiros por via marítima e fluvial
		07.3.5	Transportes combinados de passageiros
		07.3.6	Outros serviços de transportes
08			Comunicações
	08.1		Serviços postais
	08.2		Equipamentos de telecomunicações (telefónico e telefax)
	08.3		Serviços de telecomunicações
09			Lazer, recreação e cultura
	09.1		Equipamento audiovisual, fotográfico e informático
		09.1.1	Equipamento para receção, registo e reprodução de som e imagem
		09.1.2	Equipamento fotográfico, cinematográfico e instrumentos de ótica
		09.1.3	Equipamento de processamento de dados
		09.1.4	Meios ou suportes de gravação
		09.1.5	Reparação de equipamento audiovisual, fotográfico e de processamento de dados
	09.2		Outros bens duradouros de base para lazer, recreação e cultura
		09.2.1	Outros bens duradouros de base para lazer, recreação e cultura em espaços abertos
		09.2.2	Instrumentos musicais e outros bens duradouros relacionados com lazer, recreação e cultura
		09.2.3	Manutenção e reparação de outros bens duradouros relacionados com lazer, recreação e cultura
	09.3		Outros artigos e equipamento recreativos, para jardins e animais de estimação
		09.3.1	Jogos, brinquedos e artigos para atividades de recreação e lazer
		09.3.2	Equipamentos para desporto, campismo e atividades de recreação ao ar livre
		09.3.3	Jardinagem, plantas e flores
		09.3.4	Animais de estimação e produtos relacionados
		09.3.5	Serviços veterinários e outros serviços para animais de estimação
	09.4		Serviços recreativos e culturais
		09.4.1	Serviços recreativos e desportivos
		09.4.2	Serviços culturais
		09.4.3	Jogos de azar
	09.5		Jornais, livros e artigos de papelaria
		09.5.1	Livros
		09.5.2	Jornais e periódicos
		09.5.3	Impressos diversos
		09.5.4	Artigos de papelaria e de desenho
	09.6		Férias organizadas
10			Educação
		10.1	Educação pré-escolar e ensino básico (1º ciclo)

COICOP - Classificação individual do consumo dos ADP por função			
Código COICOP			Descrição
1ºNível	2ºNível	3ºNível	
		10.2	Ensino básico (2º e 3º ciclos) e ensino secundário
		10.3	Ensino pós-secundário não superior
		10.4	Ensino superior
		10.5	Ensino não definível por níveis
11			Restaurantes e hotéis
	11.1		Serviços de refeições e bebidas
		11.1.1	Restaurantes, cafés e estabelecimentos similares
		11.1.2	Cantinas ou refeitórios
	11.2		Serviços de alojamento
12			Bens e serviços diversos
	12.1		Cuidados pessoais
		12.1.1	Salões de cabeleireiro e estabelecimentos de cuidados pessoais
		12.1.2	Aparelhos elétricos para cuidados pessoais
		12.1.3	Outros aparelhos, artigos e produtos para cuidados pessoais
	12.2		Prostituição
	12.3		Artigos de uso pessoal n.e.
		12.3.1	Artigos de ourivesaria, joalheria e relojoaria
		12.3.2	Outros artigos de uso pessoal
	12.4		Proteção social
	12.5		Seguros
		12.5.1	Seguros de vida
		12.5.2	Seguros relacionados com a habitação
		12.5.3	Seguros relacionados com a saúde
		12.5.4	Seguros relacionados com os transportes
		12.5.5	Outros seguros
	12.6		Serviços financeiros n.e.
		12.6.1	Serviços financeiros sifim
		12.6.2	Outros serviços financeiros n.e.
	12.7		Outros serviços n.e.

ANEXO 2 – CALIBRAÇÃO NO SAS

O ficheiro “novo” reúne a informação do ficheiro “idef15” com a informação relativamente ao consumo realizado pelos ADP fora do território nacional, que se encontram no ficheiro “estrangeiro”. Assim, são criadas novas variáveis relativas ao consumo, realizado por classe da COICOP a primeiro nível para cada ADP, realizado unicamente em território nacional. Assim, seja $desp_{pt_j}$, a despesa realizada para cada classe da COICOP j ($\forall j=1, \dots, 12$) no território nacional para cada ADP.

Figura 1 – Comandos utilizados para a Construção do ficheiro “novo” e das novas variáveis.

```
libname novo "/sasdata/DES/CV/IDEF";

proc sort data=novo.Idef2015;
by Idagregado;
run;

proc sort data=novo.estrangeiro;
by Idagregado;
run;

data novo.novo;
merge novo.idef2015(in=x) novo.estrangeiro(in=y);
by Idagregado;
if x=1;
run;

data novo.novo;
set novo.novo;
despnt=desp_total-HJ00;
despnt_01=desp_01-HJ01;
despnt_02=desp_02-HJ02;
despnt_03=desp_03-HJ03;
despnt_04=desp_04-HJ04;
despnt_05=desp_05-HJ05;
despnt_06=desp_06-HJ06;
despnt_07=desp_07-HJ07;
despnt_08=desp_08-HJ08;
despnt_09=desp_09-HJ09;
despnt_10=desp_10-HJ10;
despnt_11=desp_11-HJ11;
despnt_12=desp_12-HJ12;
run;
```

No ficheiro “idef15” foi reunida a informação do ficheiro “novo” com os ponderadores iniciais disponíveis no ficheiro “ponderadores”. Tal foi possível através do cruzamento de informação das variáveis que identificavam os ADP em cada um dos ficheiros, o código do agregado é dado pela variável *Idagregado* para o ficheiro “idef15” e pela variável *idadp* para o ficheiro “ponderadores”. Foi também alterada a designação das variáveis *despnt_j* para *desj*.

Figura 2 – Comandos usados para a criação do ficheiro “idef15”.

```

libname idef '/sasdata/DES/CV/IDEF';

* Ficheiros enviados pelo DMSI/ME;

data idef.idef15;
set idef.idef15_nor idef.idef15_cen idef.idef15_lis idef.idef15_ale
  idef.idef15_alg idef.idef15_aco idef.idef15_mad;
run;

* Pesos finais calculados com os programas do DMSI/ME;

data idef.pesos;
set idef.pel_nor idef.pel_cen idef.pel_lis idef.pel_ale
  idef.pel_alg idef.pel_aco idef.pel_mad;
run;

* Adiciona despesa total e monetaria enviada pelo DES/CV em 22/05/2019;
* o ficheiro SPSS idef2015_adp_20190522 foi convertido para SAS atreves da opcao;
* "save as" type V9+ windows;

proc sql;
create table idef.idef15 as
select x.*, y.despnt as des, y.despmon as despm,
  y.despnt_01 as des01, y.despnt_02 as des02, y.despnt_03 as des03, y.despnt_04 as des04,
  y.despnt_05 as des05, y.despnt_06 as des06, y.despnt_07 as des07, y.despnt_08 as des08,
  y.despnt_09 as des09, y.despnt_10 as des10, y.despnt_11 as des11, y.despnt_12 as des12,
  y.despmon_01 as dem01, y.despmon_02 as dem02, y.despmon_03 as dem03, y.despmon_04 as dem04,
  y.despmon_05 as dem05, y.despmon_06 as dem06, y.despmon_07 as dem07, y.despmon_08 as dem08,
  y.despmon_09 as dem09, y.despmon_10 as dem10, y.despmon_11 as dem11, y.despmon_12 as dem12,
  z.pesofin1 as pesofin
from idef.idef15 as x left join (idef.novo as y left join idef.pesos as z on
  y.Idagregado=z.idadp)
on x.idadp=y.Idagregado;
quit;

```

A partir do “idef15”, utilizando a variável *NUTS_2*, foram criados sete ficheiros, um para cada uma das *NUTS_II* em que cada ficheiro possui apenas os ADP dessa mesma *NUTS_II*.

Figura 3 – Criação de um novo ficheiro para cada região das *NUTS_II*.

```

data idef.idef15_nor idef.idef15_cen idef.idef15_lis idef.idef15_ale
  idef.idef15_alg idef.idef15_aco idef.idef15_mad;
set idef.idef15(drop=pesofin);
if nuts_2="11" then output idef.idef15_nor;
if nuts_2="16" then output idef.idef15_cen;
if nuts_2="17" then output idef.idef15_lis;
if nuts_2="18" then output idef.idef15_ale;
if nuts_2="15" then output idef.idef15_alg;
if nuts_2="20" then output idef.idef15_aco;
if nuts_2="30" then output idef.idef15_mad;
run;

```

Na macro CALJACK o primeiro parâmetro a definir são as variáveis macro com as margens a utilizar na Calibração.

Figura 4 – Macro CALJACK: Margens

```

@MACRO CALJACK (
MARGES = {
/* Tipfa */
/*244517 421596 372453 325754
207295 309156 215915 182894
292296 381309 250160 232814
72587 109466 70003 53384
46870 64799 37778 34912
12543 22488 20346 27969
19957 25660 24526 25302*/

/* Tipin */
/*1016054 348266
457678 457582
1113020 43519
152323 153117
125210 59149
42188 41158
78406 17039*/

/* Eduref */
/*109634 833643 197218 223825
100262 556926 128219 129853
75503 539859 243657 297520
51054 174053 46304 33529
15418 104210 35627 29104
953 52844 10676 9873
8909 55851 13083 17602*/

/* Ags */
/*257853 209013 481628 494081 278060 247008 201868 513892 548863 389519
151141 118601 283786 303823 216465 143349 114507 298799 328774 304747
229478 139915 375465 338927 237617 217273 138517 405762 398919 336395
49309 36508 93592 98574 75248 46902 35001 92809 101242 104185
34455 22880 59066 56558 39797 32714 21943 62170 60579 51306
21170 17328 38298 30834 13128 20311 16704 37562 31536 19482
20080 16853 37740 32527 13868 19198 15843 38612 38615 25350*/

/* Edu */
/*379046 2179819 573241 489679
257577 1309656 379669 317090
282626 1359203 576097 591242
107031 414301 124813 87225
32470 234714 94700 59584
25080 156350 38395 26528
31691 150313 44235 32447
*/

/* COICOP 01-12 NUTS 2*/

/*6919406243 1251208026 2607774283 7683399172 2067256194 2069817509 4984623778 963201094 2465706428 482836333 4754150693 4240969466
4641913743 839378341 1749436709 5154441719 1386827800 1388546069 3343956520 646167639 1654129871 323912852 3189342649 2845072794
5865605706 10606533570 2210619694 6513245276 1752420555 1754591791 4225483611 816509052 2090188259 409302106 4030110743 3595085160
1549096578 280116820 583820934 1720137096 462811314 463384734 1115943097 215638664 552015195 108095996 1064345452 949455930
935011446 169074308 352385554 1038248936 279345963 279692071 673566506 130156261 333180087 65245121 642422942 573077350
422704961 76436015 159308341 469377116 126288213 126444683 304509538 58841737 150629448 29496362 290429990 259079865
484067322 87531921 182434485 537514684 144620960 144800144 348713950 67383553 172495712 33778229 332590531 296689436
*/
},
);

```

Descrição das variáveis utilizadas como margens para a calibração:

- Tipfa é a variável que representa o total de ADP em cada tipo de família, considerando as 4 seguintes categorias: 1 pessoa, 2 pessoas, 3 pessoas, 4 e mais pessoas.
- Tipin é a variável que representa o total de ADP em tipo de aglomeração, urbano ou não urbano.
- A variável Eduref é a variável que representa o total de ADP em cada um dos quatro níveis de escolaridade do indivíduo de referência do agregado (nenhum, ensino básico, ensino secundário, ensino superior).
- Ags é a variável que representa o total de indivíduos em cada um dos cinco escalões etários (0-14 anos, 15-24 anos, 25-44 anos, 45-64 anos, 65 e mais anos) por sexo.

- Edu representa o total de indivíduos em cada um dos níveis de escolaridade (nenhum, ensino básico, ensino secundário, ensino superior).
- *COICOP 01-12 NUTS_II*, representa o consumo por primeiro nível da COICOP, por NUTS_II segundo as Contas Nacionais.

ANEXO 3 – VARIÁVEIS MACRO

Algumas designações das variáveis macro estão em francês. A macro utiliza um conjunto de variáveis cuja definição é apresentada de seguida.

- MARGENS - Variável macro que contém as margens utilizadas na calibração, recorde-se que foram usadas as margens utilizadas pelo IDEF 2015/2016 e a nova margem criada com a informação auxiliar das Contas Nacionais (veja-se anexo 2 – figura 4). Dado as CN apenas possuem informação a nível nacional foi realizada a desagregação por NUTS_II e COICOP que é proporcional ao número de ADP por NUTS_II, utilizando informação obtida através do próprio IDEF.
- DATA – Variável macro que contém o nome do ficheiro SAS de entrada, com a informação dos ponderadores iniciais e das variáveis desenvolvidas no IDEF não calibradas para a NUTS_II respetiva. Assim, por exemplo, no caso da região norte o ficheiro de entrada será “idef15_nor”.
- M – Variável macro que indica a função distância a usar identificada por um número inteiro de 1 a 7, correspondendo ao método logit o número 3. Assim “M”=3, para qualquer região.
- VQUAL – Variável macro que contém a lista de nomes das variáveis qualitativas utilizadas na calibração.
- VQUAN – Variável macro que contém a lista de nomes das variáveis quantitativas utilizadas na calibração.
- NBMOD – Variável macro que indica o número de modalidades das variáveis qualitativas, separadas por espaços brancos e pela mesma ordem da lista VQUAL.
- POIDS – Variável macro que indica o nome da variável que contém os ponderadores iniciais dos ADP da amostra antes da extrapolação (“pesoin”).
- LO – Limite inferior (L) da função logit.

- UP – Limite superior (U) da função logit.
- STRATE² – Variável macro que contém o nome da variável que identifica o estrato (STRATE = “NUTS_2”).
- REPLIC² – Variável macro que contém o nome da variável que identifica as réplicas (REPLIC = “rep”), para o cálculo da variância de Jackknife.
- MENAGE² – Variável macro que contém o nome da variável que identifica os ADP (MENAGE = “idadp”³).
- TRANSF: Se TRANS=O (*Oui* em francês) é realizada uma transformação para que todos os membros do agregado tenham o mesmo peso final. Como no IDEF essa questão não se levanta colocamos TRANSF = N (*Non* em francês).
- CALMAR: Se CALMAR=O (*Oui* em francês) permite guardar os ponderadores ajustados e calcular as variâncias calibradas desde que CAL_EST=O.
- NOM_PDS – Variável macro que define o nome que se pretende para a variável de ponderação final, ou seja, para o peso ajustado (NOM_PDS = “pesofin3”)
- FICH_PDS – Variável macro que contém o nome do ficheiro SAS criado com os ponderadores finais (FICH_PDS = “pe3_nor” para o caso da região norte)
- CALC_LAM: Se CALC_LAM=O (*Oui* em francês) para que possam ser calculados e guardados os multiplicadores de Lagrange (lambdas).
- FICH_LAM – Parâmetro que contém o nome do ficheiro SAS criado com os lambdas (FICH_LAM = “la3_nor” para o caso da região norte)
- CALC_EST – Variável macro que permite o cálculo das estimativas do total para as variáveis de estudo pretendidas CALC_EST=O (*Oui* em francês). Para realizar este cálculo: CALMAR=O e CALC_LAM=O.
- FICH_EST – Variável macro que contém o nome do ficheiro SAS criado com as variáveis calibradas (FICH_EST = “es3_nor” para o caso da região norte).

² Variável nomeada em Francês.

³ “idadp” – variável que identifica os ADP no ficheiro de entrada.

- CALC_VAR: Se CALC_VAR=O (*Oui* em francês) é calculado o estimador para a variância de Jackknife para a variância de cada variável calibrada.
- FICH_VAR - Parâmetro que contém o nome do ficheiro SAS criado com as variâncias (FICH_VAR = “var3_nor” para o caso da região norte).
- FICH_CV – Variável macro que contém o nome do ficheiro SAS criado com os coeficientes de variação (FICH_CV = “cv3_nor” para o caso da região norte).
- VETUDE – Variável macro que contém os nomes das variáveis de estudo utilizadas para estimar os totais e as respetivas variâncias.

Para além de ser realizada a calibração para todas as variáveis definidas em VQUAL e VQUAN, é ainda realizada a calibração para o rácio correspondente à despesa, média por ADP realizada em cada COICOP ou agrupamento de divisões da COICOP na região para a qual é realizada a calibração. Para tal é necessário então definir os numeradores e os denominadores. Assim o número de variáveis definidas em NUMER tem de ser igual ao número de variáveis definidas em DENOM.

- NUMER – Parâmetro que contém o nome das variáveis de estudo para os numeradores utilizados nas estimativas de quocientes e das suas variâncias.
- DENOM - Parâmetro que contém o nome das variáveis de estudo para os denominadores utilizados nas estimativas de quocientes e das suas variâncias. A variável utilizada como denominador é a variável v1 que representa o total de ADP para cada NUTS_II em análise.

ANEXO 4 – CALIBRAÇÃO PARA CADA REGIÃO DA NUTS_II.

Figura 5 – Macro CALJACK: Norte

```

DATA      =  idef.idef15_nor,          /* TABLE SAS EN ENTREE */
M         =  3,                      /* FONCTION DISTANCE*/
VQUAL    =  tipfa tipin edure,       /* VARIABLES QUALITATIVES */
VQUAN    =  ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10 /* VARIABLES QUANTITATIVES */
          edu01 edu02 edu03 /*edu04*/
          des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
NEMOD    =  4 2 4,                  /* NOMBRE DE MODALITES DES VARIABLES QUALITATIVES */
POIDS    =  pesoin,                 /* PONDERATION INITIALE (SONDAGE) */
LO       =  0.25,                   /* BORNE INF (LOGIT OU LINEAIRE TRONQUEE)*/
UP       =  4.00,                   /* BORNE SUP (LOGIT OU LINEAIRE TRONQUEE)*/
STRATE   =  nuts_2,                 /* IDENTIFIE LA STRATE */
REPLIC   =  rep,                    /* IDENTIFIE LA REPLIQUE */
MENAGE   =  idadp,                  /* CONTIENT LES VARIABLES IDENTIFIANT LES MENAGES */
TRANSF   =  N,                      /* MEME POIDS FINAL AU MENAGE */
CALMAR   =  O,                      /* CALCUL ET SAUVEGARDE DES POIDS FINAUX */
          /* CALJACK JOUE LE MEME ROLE QUE CALMAR */
NOM_PDS  =  pesofin3,               /* NOM DE LA VARIABLE DE PONDERATION FINALE */
FICH_PDS =  idef.pe3_nor,           /* NOM DU FICHIER DES POIDS SI &CALMAR=O */
CALC_LAM =  O,                      /* CALCUL DES LAMBDA */
FICH_LAM =  idef.la3_nor,           /* NOM DU FICHIER DES LAMBDA */
CALC_EST =  O,                      /* CALCUL D'ESTIMATIONS */
FICH_EST =  idef.es3_nor,           /* NOM DU FICHIER DES ESTIMATIONS */
CALC_VAR =  O,                      /* CALCUL ESTIMATION JACKKNIFE DE VARIANCE */
FICH_VAR =  idef.va3_nor,           /* NOM DU FICHIER DES VARIANCES */
FICH_CV  =  idef.cv3_nor,           /* NOM DU FICHIER DES COEFF. DE VARIATION */
VETUDE   =  v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11
          ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10
          edu01 edu02 edu03 edu04
          des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
DIFF_TOT =  ,                       /* PAIRES DE TOTAUX POUR DIFFERENCES */
NUMER    =  des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
DENOM    =  v1 v1,

```

Figura 6 – Macro CALJACK: Centro

```

DATA      =  idef.idef15_cen,        /* TABLE SAS EN ENTREE */
M         =  3,                      /* FONCTION DISTANCE*/
VQUAL    =  tipfa tipin edure,       /* VARIABLES QUALITATIVES */
VQUAN    =  ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10 /* VARIABLES QUANTITATIVES */
          edu01 edu02 edu03 /*edu04*/
          des01 /*des02*/ des03 des04 /*des05 des06*/ des07 /*des08 des09 des10*/ des11 des12 des16 des17 des18,
NEMOD    =  4 2 4,                  /* NOMBRE DE MODALITES DES VARIABLES QUALITATIVES */
POIDS    =  pesoin,                 /* PONDERATION INITIALE (SONDAGE) */
LO       =  0.04,                   /* BORNE INF (LOGIT OU LINEAIRE TRONQUEE)*/
UP       =  25.00,                  /* BORNE SUP (LOGIT OU LINEAIRE TRONQUEE)*/
STRATE   =  nuts_2,                 /* IDENTIFIE LA STRATE */
REPLIC   =  rep,                    /* IDENTIFIE LA REPLIQUE */
MENAGE   =  idadp,                  /* CONTIENT LES VARIABLES IDENTIFIANT LES MENAGES */
TRANSF   =  N,                      /* MEME POIDS FINAL AU MENAGE */
CALMAR   =  O,                      /* CALCUL ET SAUVEGARDE DES POIDS FINAUX */
          /* CALJACK JOUE LE MEME ROLE QUE CALMAR */
NOM_PDS  =  pesofin3,               /* NOM DE LA VARIABLE DE PONDERATION FINALE */
FICH_PDS =  idef.pe3_cen,           /* NOM DU FICHIER DES POIDS SI &CALMAR=O */
CALC_LAM =  O,                      /* CALCUL DES LAMBDA */
FICH_LAM =  idef.la3_cen,           /* NOM DU FICHIER DES LAMBDA */
CALC_EST =  O,                      /* CALCUL D'ESTIMATIONS */
FICH_EST =  idef.es3_cen,           /* NOM DU FICHIER DES ESTIMATIONS */
CALC_VAR =  O,                      /* CALCUL ESTIMATION JACKKNIFE DE VARIANCE */
FICH_VAR =  idef.va3_cen,           /* NOM DU FICHIER DES VARIANCES */
FICH_CV  =  idef.cv3_cen,           /* NOM DU FICHIER DES COEFF. DE VARIATION */
VETUDE   =  v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11
          ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10
          edu01 edu02 edu03 edu04
          des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des16 des17 des18,
DIFF_TOT =  ,                       /* PAIRES DE TOTAUX POUR DIFFERENCES */
NUMER    =  des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des16 des17 des18,
DENOM    =  v1 v1,

```

Figura 7 – Criação das variáveis *des16*, *des17* e *des18* (Centro)

```

/*Centro*/
data ndef.ndef15_cen;
set ndef.ndef15_cen;
des16=des09+des06;
run;

data ndef.ndef15_cen;
set ndef.ndef15_cen;
des17=des05+des02;
run;

data ndef.ndef15_cen;
set ndef.ndef15_cen;
des18=des08+des10;
run;

```

Nota: A variável *des16* corresponde à soma da variável *des09* com a variável *des06*; A variável *des17* corresponde à soma da variável *des05* com a variável *des02*; A variável *des18* corresponde à soma da variável *des08* com a variável *des10*;

Figura 8 – Macro CALJACK: Lisboa

```

DATA      = ndef.ndef15_lis,          /* TABLE SAS EN ENTREE */
M         = 3,                       /* FONCTION DISTANCE*/
VQUAL    = tipfa tipin edure,        /* VARIABLES QUALITATIVES */
VQUAN    = ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10 /* VARIABLES QUANTITATIVES */
          edu01 edu02 edu03 /*edu04*/
          des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
NEMOD    = 4 2 4,                   /* NOMBRE DE MODALITES DES VARIABLES QUALITATIVES */
POIDS    = pesoin,                  /* PONDERATION INITIALE (SONDAGE) */
LO       = 0.045,                   /* BORNE INF (LOGIT OU LINEAIRE TRONQUEE)*/
UP       = 22.00,                   /* BORNE SUP (LOGIT OU LINEAIRE TRONQUEE)*/
STRATE   = nuts_2,                  /* IDENTIFIE LA STRATE */
REPLIC   = rep,                     /* IDENTIFIE LA REPLIQUE */
MENAGE   = idadp,                   /* CONTIENT LES VARIABLES IDENTIFIANT LES MENAGES */
TRANSF   = N,                       /* MEME POIDS FINAL AU MENAGE */
CALMAR   = 0,                       /* CALCUL ET SAUVEGARDE DES POIDS FINAUX */
          /* CALJACK JOUE LE MEME ROLE QUE CALMAR */
NOM_PDS  = pesofin3,                /* NOM DE LA VARIABLE DE PONDERATION FINALE */
FICH_PDS = ndef.pe3_lis,            /* NOM DU FICHIER DES POIDS SI &CALMAR=0 */
CALC_LAM = 0,                       /* CALCUL DES LAMBDA */
FICH_LAM = ndef.la3_lis,            /* NOM DU FICHIER DES LAMBDA */
CALC_EST = 0,                       /* CALCUL D'ESTIMATIONS */
FICH_EST = ndef.es3_lis,            /* NOM DU FICHIER DES ESTIMATIONS */
CALC_VAR = 0,                       /* CALCUL ESTIMATION JACKKNIFE DE VARIANCE */
FICH_VAR = ndef.va3_lis,            /* NOM DU FICHIER DES VARIANCES */
FICH_CV  = ndef.cv3_lis,            /* NOM DU FICHIER DES COEFF. DE VARIATION */
VETUDE   = v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11
          ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10
          edu01 edu02 edu03 edu04
          des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
DIFF_TOT = ,                        /* PAIRES DE TOTAUX POUR DIFFERENCES */
NUMER    = des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
DENOM    = v1 v1,
DIF_RAT  = ,                        /* PAIRES DE RATIOS POUR DIFFERENCES */
MAXITER  = 20,                      /* NOMBRE D'ITERATIONS MAXIMALES PERMISES */
SEUIL    = 0.0005;                  /* SEUIL POUR LE TEST D'ARRET */

```

Figura 9 – Macro CALJACK: Alentejo

```

DATA      = idef.idef15_ale,          /* TABLE SAS EN ENTREE */
M         = 3,                      /* FONCTION DISTANCE*/
VQUAL    = tipfa tipin edure,       /* VARIABLES QUALITATIVES */
VQUAN    = ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10 /* VARIABLES QUANTITATIVES */
          edu01 edu02 edu03 /*edu04*/
          des01 /*des02 des03*/ des04 /*des05 des06 des07 des08 des09 des10 des11 des12*/ des26,
NBMOD    = 4 2 4,                  /* NOMBRE DE MODALITES DES VARIABLES QUALITATIVES */
POIDS    = pesocin,                /* PONDERATION INITIALE (SONDAGE) */
LO       = 0.03,                  /* BORNE INF (LOGIT OU LINEAIRE TRONQUEE)*/
UP       = 30.00,                 /* BORNE SUP (LOGIT OU LINEAIRE TRONQUEE)*/
STRATE   = nuts_2,                /* IDENTIFIE LA STRATE */
REPLIC   = rep,                   /* IDENTIFIE LA REPLIQUE */
MENAGE   = idadp,                 /* CONTIENT LES VARIABLES IDENTIFIANT LES MENAGES */
TRANSF   = N,                     /* MEME POIDS FINAL AU MENAGE */
CALMAR   = O,                     /* CALCUL ET SAUVEGARDE DES POIDS FINAUX */
          /* CALJACK JOUE LE MEME ROLE QUE CALMAR */
NOM_PDS  = pesofin3,              /* NOM DE LA VARIABLE DE PONDERATION FINALE */
FICH_PDS = idef.pe3_ale,          /* NOM DU FICHIER DES POIDS SI &CALMAR=O */
CALC_LAM = O,                     /* CALCUL DES LAMBDA */
FICH_LAM = idef.la3_ale,          /* NOM DU FICHIER DES LAMBDA */
CALC_EST = O,                     /* CALCUL D'ESTIMATIONS */
FICH_EST = idef.es3_ale,          /* NOM DU FICHIER DES ESTIMATIONS */
CALC_VAR = O,                     /* CALCUL ESTIMATION JACKKNIFE DE VARIANCE */
FICH_VAR = idef.va3_ale,          /* NOM DU FICHIER DES VARIANCES */
FICH_CV  = idef.cv3_ale,          /* NOM DU FICHIER DES COEFF. DE VARIATION */
VETUDE   = v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11
          ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10
          edu01 edu02 edu03 edu04
          des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des26,
DIFF_TOT = ,                       /* PAIRES DE TOTAUX POUR DIFFERENCES */
NUMER    = des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des26,
DENOM    = v1 v1,

```

Figura 10 – Criação da variável *des26* (Alentejo)

```

/*Alentejo*/
data idef.idef15_ale;
set idef.idef15_ale;
des26=des07+des11+des12+des03+des09+des06+des05+des02+des08+des10;
run;

```

Nota: A variável *des26* corresponde à soma das variáveis *des07*, *des11*, *des12*, *des03*, *des09*, *des06*, *des05*, *des02*, *des08*, *des10*;

Figura 11 – Macro CALJACK: Algarve

```

DATA      =  idef.idef15_alg,          /* TABLE SAS EN ENTREE */
M         =  3,                      /* FONCTION DISTANCE*/
VQUAL    =  tipfa tipin edure,       /* VARIABLES QUALITATIVES */
VQUAN    =  ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10 /* VARIABLES QUANTITATIVES */
          edu01 edu02 edu03 /*edu04*/
          des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
NBMOD    =  4 2 4,                  /* NOMBRE DE MODALITES DES VARIABLES QUALITATIVES */
POIDS    =  pescoin,                /* PONDERATION INITIALE (SONDAGE) */
LO       =  0.04,                  /* BORNE INF (LOGIT OU LINEAIRE TRONQUEE)*/
UP       =  25.00,                 /* BORNE SUP (LOGIT OU LINEAIRE TRONQUEE)*/
STRATE   =  nuts_2,                /* IDENTIFIE LA STRATE */
REPLIC   =  rep,                  /* IDENTIFIE LA REPLIQUE */
MENAGE   =  idadp,                /* CONTIENT LES VARIABLES IDENTIFIANT LES MENAGES */
TRANSF   =  N,                    /* MEME POIDS FINAL AU MENAGE */
CALMAR   =  O,                    /* CALCUL ET SAUVEGARDE DES POIDS FINAUX */
          /* CALJACK JOUE LE MEME ROLE QUE CALMAR */
NOM_PDS  =  pesofin3,              /* NOM DE LA VARIABLE DE PONDERATION FINALE */
FICH_PDS =  idef.pe3_alg,          /* NOM DU FICHIER DES POIDS SI &CALMAR=O */
CALC_LAM =  O,                    /* CALCUL DES LAMBDA */
FICH_LAM =  idef.la3_alg,          /* NOM DU FICHIER DES LAMBDA */
CALC_EST =  O,                    /* CALCUL D'ESTIMATIONS */
FICH_EST =  idef.es3_alg,          /* NOM DU FICHIER DES ESTIMATIONS */
CALC_VAR =  O,                    /* CALCUL ESTIMATION JACKKNIFE DE VARIANCE */
FICH_VAR =  idef.va3_alg,          /* NOM DU FICHIER DES VARIANCES */
FICH_CV  =  idef.cv3_alg,          /* NOM DU FICHIER DES COEFF. DE VARIATION */
VETUDE   =  v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11
          ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10
          edu01 edu02 edu03 edu04
          des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
DIFF_TOT = ,                      /* PAIRES DE TOTAUX POUR DIFFERENCES */
NUMER    =  des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12,
DENOM    =  v1 v1,

```

Figura 12 – Macro CALJACK: Região Autónoma dos Açores

```

DATA      =  idef.idef15_aco,        /* TABLE SAS EN ENTREE */
M         =  3,                      /* FONCTION DISTANCE*/
VQUAL    =  tipfa tipin edure,       /* VARIABLES QUALITATIVES */
VQUAN    =  ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10 /* VARIABLES QUANTITATIVES */
          edu01 edu02 edu03 /*edu04*/
          des01 /*des02 des03*/ des04 /*des05 des06 des07 des08 des09 des10 des11 des12*/ des27,
NBMOD    =  4 2 4,                  /* NOMBRE DE MODALITES DES VARIABLES QUALITATIVES */
POIDS    =  pescoin,                /* PONDERATION INITIALE (SONDAGE) */
LO       =  0.03,                  /* BORNE INF (LOGIT OU LINEAIRE TRONQUEE)*/
UP       =  30.00,                 /* BORNE SUP (LOGIT OU LINEAIRE TRONQUEE)*/
STRATE   =  nuts_2,                /* IDENTIFIE LA STRATE */
REPLIC   =  rep,                  /* IDENTIFIE LA REPLIQUE */
MENAGE   =  idadp,                /* CONTIENT LES VARIABLES IDENTIFIANT LES MENAGES */
TRANSF   =  N,                    /* MEME POIDS FINAL AU MENAGE */
CALMAR   =  O,                    /* CALCUL ET SAUVEGARDE DES POIDS FINAUX */
          /* CALJACK JOUE LE MEME ROLE QUE CALMAR */
NOM_PDS  =  pesofin3,              /* NOM DE LA VARIABLE DE PONDERATION FINALE */
FICH_PDS =  idef.pe3_aco,          /* NOM DU FICHIER DES POIDS SI &CALMAR=O */
CALC_LAM =  O,                    /* CALCUL DES LAMBDA */
FICH_LAM =  idef.la3_aco,          /* NOM DU FICHIER DES LAMBDA */
CALC_EST =  O,                    /* CALCUL D'ESTIMATIONS */
FICH_EST =  idef.es3_aco,          /* NOM DU FICHIER DES ESTIMATIONS */
CALC_VAR =  O,                    /* CALCUL ESTIMATION JACKKNIFE DE VARIANCE */
FICH_VAR =  idef.va3_aco,          /* NOM DU FICHIER DES VARIANCES */
FICH_CV  =  idef.cv3_aco,          /* NOM DU FICHIER DES COEFF. DE VARIATION */
VETUDE   =  v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11
          ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10
          edu01 edu02 edu03 edu04
          des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des27,
DIFF_TOT = ,                      /* PAIRES DE TOTAUX POUR DIFFERENCES */
NUMER    =  des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des27,
DENOM    =  v1 v1,

```

Figura 13 – Criação da variável *des27* (Açores)

```

/*Açores*/
data ndef.ndef15_aco;
set ndef.ndef15_aco;
des27=des07+des11+des12+des03+des09+des06+des05+des02+des08+des10;
run;

```

Nota: A variável *des26* corresponde à soma das variáveis *des07*, *des11*, *des12*, *des03*, *des09*, *des06*, *des05*, *des02*, *des08*, *des10*;

Figura 14 – Macro CALJACK: Região Autónoma da Madeira

```

DATA      = ndef.ndef15_mad,          /* TABLE SAS EN ENTREE */
M         = 3,                       /* FONCTION DISTANCE*/
VQUAL    = tipfa tipin edure,        /* VARIABLES QUALITATIVES */
VQUAN    = ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10 /* VARIABLES QUANTITATIVES */
          edu01 edu02 edu03 /*edu04*/
          /*des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12*/ des13 des14 des23 des24 des25 des26,
NEMOD    = 4 2 4,                   /* NOMBRE DE MODALITES DES VARIABLES QUALITATIVES */
POIDS    = pesoin,                  /* PONDERATION INITIALE (SONDAGE) */
LO       = 0.07,                    /* BORNE INF (LOGIT OU LINEAIRE TRONQUEE)*/
UP       = 15.00,                   /* BORNE SUP (LOGIT OU LINEAIRE TRONQUEE)*/
STRATE   = nuts_2,                  /* IDENTIFIE LA STRATE */
REPLIC   = rep,                     /* IDENTIFIE LA REPLIQUE */
MENAGE   = idadp,                   /* CONTIENT LES VARIABLES IDENTIFIANT LES MENAGES */
TRANSF   = N,                       /* MEME POIDS FINAL AU MENAGE */
CALMAR   = O,                       /* CALCUL ET SAUVEGARDE DES POIDS FINAUX */
          /* CALJACK JOUE LE MEME ROLE QUE CALMAR */
NOM_PDS  = pesofin3,                /* NOM DE LA VARIABLE DE PONDERATION FINALE */
FICH_PDS = ndef.pe3_mad,             /* NOM DU FICHIER DES POIDS SI &CALMAR=O */
CALC_LAM = O,                       /* CALCUL DES LAMBDA */
FICH_LAM = ndef.la3_mad,            /* NOM DU FICHIER DES LAMBDA */
CALC_EST = O,                       /* CALCUL D'ESTIMATIONS */
FICH_EST = ndef.es3_mad,            /* NOM DU FICHIER DES ESTIMATIONS */
CALC_VAR = O,                       /* CALCUL ESTIMATION JACKKNIFE DE VARIANCE */
FICH_VAR = ndef.va3_mad,            /* NOM DU FICHIER DES VARIANCES */
FICH_CV  = ndef.cv3_mad,            /* NOM DU FICHIER DES COEFF. DE VARIATION */
VETUDE   = v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11
          ags01 ags02 ags03 ags04 ags05 ags06 ags07 ags08 ags09 ags10
          edu01 edu02 edu03 edu04
          des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des13 des14 des23 des24 des25 des26,
DIFF_TOT = ,                         /* PAIRES DE TOTAUX POUR DIFFERENCES */
NUMER    = des des01 des02 des03 des04 des05 des06 des07 des08 des09 des10 des11 des12 des13 des14 des23 des24 des25 des26,
DENOM    = v1 v1,

```

Figura 15 – Criação das variáveis *des13*, *des14*, *des23*, *des25* (Madeira)

```

/*Madeira*/
data ndef.ndef15_mad;
set ndef.ndef15_mad;
des13=des04+des01;
run;

data ndef.ndef15_mad;
set ndef.ndef15_mad;
des14=des07+des11;
run;

data ndef.ndef15_mad;
set ndef.ndef15_mad;
des23=des12+des03;
run;

data ndef.ndef15_mad;
set ndef.ndef15_mad;
des24=des09+des06;
run;

data ndef.ndef15_mad;
set ndef.ndef15_mad;
des25=des05+des02;
run;

```

Nota: A variável *des13* corresponde à soma da variável *des04* com a variável *des01*; A variável *des14* corresponde à soma da variável *des07* com a variável *des11*; A variável *des23* corresponde à soma da variável *des12* com a variável *des03*; A variável *des24* corresponde à soma da variável *des09* com a variável *des06*; A variável *des25* corresponde à soma da variável *des05* com a variável *des02*;

ANEXO 5 – COEFICIENTE DE VARIAÇÃO

Tabela 6 - Comparação do Coeficiente de Variação para a região Norte

Descrição da variável	Calibração IDEF2015	Calibração com as CN
Rácio do total da despesa por ADP	1,524	1,30E-08
Rácio da despesa na COICOP 01 por ADP	1,899	1,45E-09
Rácio da despesa na COICOP 02 por ADP	5,032	3,98E-09
Rácio da despesa na COICOP 03 por ADP	3,991	9,75E-09
Rácio da despesa na COICOP 04 por ADP	1,529	1,62E-08
Rácio da despesa na COICOP 05 por ADP	4,328	3,02E-08
Rácio da despesa na COICOP 06 por ADP	2,844	4,80E-08
Rácio da despesa na COICOP 07 por ADP	3,569	1,94E-08
Rácio da despesa na COICOP 08 por ADP	1,412	8,33E-09
Rácio da despesa na COICOP 09 por ADP	3,600	2,98E-08
Rácio da despesa na COICOP 10 por ADP	8,783	2,04E-08

Rácio da despesa na COICOP 11 por ADP	4,002	7,24E-09
Rácio da despesa na COICOP 12 por ADP	3,137	1,47E-08

Tabela 7 - Comparação do Coeficiente de Variação para o Centro

Descrição da variável	Calibração IDEF2015	Calibração com as CN
Rácio do total da despesa por ADP	1,997	6,57E-02
Rácio da despesa na COICOP 01 por ADP	2,542	3,45E-10
Rácio da despesa na COICOP 02 por ADP	7,672	-
Rácio da despesa na COICOP 03 por ADP	4,623	2,98E-08
Rácio da despesa na COICOP 04 por ADP	1,535	7,84E-09
Rácio da despesa na COICOP 05 por ADP	6,420	-
Rácio da despesa na COICOP 06 por ADP	3,393	-
Rácio da despesa na COICOP 07 por ADP	5,522	5,32E-09
Rácio da despesa na COICOP 08 por ADP	1,671	-
Rácio da despesa na COICOP 09 por ADP	4,813	-
Rácio da despesa na COICOP 10 por ADP	7,706	-
Rácio da despesa na COICOP 11 por ADP	6,576	1,17E-08
Rácio da despesa na COICOP 12 por ADP	2,901	1,79E-08
Rácio da variável "des16" por ADP	-	2,08E-09
Rácio da variável "des17" por ADP	-	1,66E-08
Rácio da variável "des18" por ADP	-	4,90E-08

Tabela 8 - Comparação do Coeficiente de Variação para Lisboa

Descrição da variável	Calibração IDEF2015	Calibração com as CN
Rácio do total da despesa por ADP	1,533	6,4E-09
Rácio da despesa na COICOP 01 por ADP	1,827	1,3E-08
Rácio da despesa na COICOP 02 por ADP	4,932	5,7E-09
Rácio da despesa na COICOP 03 por ADP	3,634	4E-08
Rácio da despesa na COICOP 04 por ADP	1,342	7,8E-09
Rácio da despesa na COICOP 05 por ADP	5,411	3,1E-08
Rácio da despesa na COICOP 06 por ADP	2,306	2,3E-08
Rácio da despesa na COICOP 07 por ADP	3,287	1,1E-08
Rácio da despesa na COICOP 08 por ADP	1,489	4,3E-09
Rácio da despesa na COICOP 09 por ADP	3,927	1,1E-08
Rácio da despesa na COICOP 10 por ADP	6,101	1,8E-08
Rácio da despesa na COICOP 11 por ADP	3,708	1,8E-08
Rácio da despesa na COICOP 12 por ADP	2,742	1,8E-08

Tabela 9 - Comparação do Coeficiente de Variação para o Alentejo

Descrição da variável	Calibração IDEF2015	Calibração com as CN
Rácio do total da despesa por ADP	2,394	0,5426
Rácio da despesa na COICOP 01 por ADP	3,778	0,13167
Rácio da despesa na COICOP 02 por ADP	8,477	-
Rácio da despesa na COICOP 03 por ADP	5,661	-
Rácio da despesa na COICOP 04 por ADP	1,942	0,75918
Rácio da despesa na COICOP 05 por ADP	6,853	-
Rácio da despesa na COICOP 06 por ADP	4,510	-
Rácio da despesa na COICOP 07 por ADP	9,538	-
Rácio da despesa na COICOP 08 por ADP	1,540	-
Rácio da despesa na COICOP 09 por ADP	6,183	-
Rácio da despesa na COICOP 10 por ADP	12,325	-
Rácio da despesa na COICOP 11 por ADP	6,447	-
Rácio da despesa na COICOP 12 por ADP	4,231	-
Rácio da variável "des26" por ADP	-	10,2184

Tabela 10 - Comparação do Coeficiente de Variação para o Algarve

Descrição da variável	Calibração IDEF2015	Calibração com as CN
Rácio do total da despesa por ADP	2,143	7,4E-09
Rácio da despesa na COICOP 01 por ADP	3,249	1,8E-10
Rácio da despesa na COICOP 02 por ADP	6,691	2,7E-09
Rácio da despesa na COICOP 03 por ADP	5,538	6,7E-09
Rácio da despesa na COICOP 04 por ADP	1,993	1,4E-09
Rácio da despesa na COICOP 05 por ADP	5,845	1,8E-08
Rácio da despesa na COICOP 06 por ADP	3,733	9,2E-09
Rácio da despesa na COICOP 07 por ADP	6,152	8,7E-09
Rácio da despesa na COICOP 08 por ADP	2,100	3,7E-11
Rácio da despesa na COICOP 09 por ADP	6,435	1,2E-08
Rácio da despesa na COICOP 10 por ADP	9,511	1,1E-08
Rácio da despesa na COICOP 11 por ADP	5,286	1,1E-08
Rácio da despesa na COICOP 12 por ADP	5,417	1,2E-08

Tabela 11 - Comparação do Coeficiente de Variação para a Região Autónoma dos Açores

Descrição da variável	Calibração IDEF2015	Calibração com as CN
Rácio do total da despesa por ADP	2,061	9,3E-09
Rácio da despesa na COICOP 01 por ADP	3,374	5,2E-09
Rácio da despesa na COICOP 02 por ADP	8,205	-
Rácio da despesa na COICOP 03 por ADP	5,491	-

Rácio da despesa na COICOP 04 por ADP	1,621	7,8E-09
Rácio da despesa na COICOP 05 por ADP	6,760	-
Rácio da despesa na COICOP 06 por ADP	4,082	-
Rácio da despesa na COICOP 07 por ADP	5,370	-
Rácio da despesa na COICOP 08 por ADP	1,368	-
Rácio da despesa na COICOP 09 por ADP	5,366	-
Rácio da despesa na COICOP 10 por ADP	11,344	-
Rácio da despesa na COICOP 11 por ADP	7,583	-
Rácio da despesa na COICOP 12 por ADP	3,972	-
Rácio da variável "des27" por ADP	-	1,5E-08

Tabela 12 - Comparação do Coeficiente de Variação para a Região Autónoma da Madeira

Descrição da variável	Calibração IDEF2015	Calibração com as CN
Rácio do total da despesa por ADP	2,213	1,4E-08
Rácio da despesa na COICOP 01 por ADP	2,396	-
Rácio da despesa na COICOP 02 por ADP	7,186	-
Rácio da despesa na COICOP 03 por ADP	6,935	-
Rácio da despesa na COICOP 04 por ADP	1,640	-
Rácio da despesa na COICOP 05 por ADP	6,429	-
Rácio da despesa na COICOP 06 por ADP	4,287	-
Rácio da despesa na COICOP 07 por ADP	6,120	-
Rácio da despesa na COICOP 08 por ADP	1,651	-
Rácio da despesa na COICOP 09 por ADP	6,245	-
Rácio da despesa na COICOP 10 por ADP	6,465	-
Rácio da despesa na COICOP 11 por ADP	6,526	-
Rácio da despesa na COICOP 12 por ADP	4,736	-
Rácio da variável "des13" por ADP	-	2,3E-08
Rácio da variável "des14" por ADP	-	3,3E-09
Rácio da variável "des23" por ADP	-	6,6E-09
Rácio da variável "des24" por ADP	-	2,9E-09
Rácio da variável "des25" por ADP	-	1,5E-08
Rácio da variável "des26" por ADP	-	3,8E-08