

“Investigação Científica e Segredo Estatístico”

Comunicação a apresentar no *Seminário sobre o Princípio do Segredo Estatístico*
organizado por:

- Conselho Superior de Estatística
- Gabinete de Política Legislativa e Planeamento do Ministério da Justiça
- Instituto Nacional de Estatística

Natércia dos Anjos Arranhado Silveira Godinho Mira

Departamento de Economia

Universidade de Évora

Email: nmira@uevora.pt

Lisboa, 13 de Janeiro de 2005

“Investigação Científica e Segredo Estatístico”

Nota Prévia

Esta comunicação é o resultado de uma reflexão efectuada sobre os limites à investigação científica que até à data e à luz da legislação em vigor, quer em Portugal, quer na Comunidade, têm sido “impostos” pelo segredo estatístico. Essa reflexão, achámos por bem complementá-la com aquela que foi a experiência de três investigadores da Universidade de Évora, que pela primeira vez solicitaram ao INE a libertação do segredo estatístico relativamente a microdados. Foi na sequência desse processo de pedido de libertação do segredo, que nos foi endereçado pela Organização o convite para participação neste seminário subordinado ao tema “Investigação Científica e Segredo Estatístico”, convite esse que desde já agradecemos.

1. A informação e o segredo estatístico: enquadramento

Não existe margem para dúvidas, que na actual sociedade da informação e do conhecimento, a informação em geral se revela um recurso crucial para a actividade humana, na medida em que é através dela que se consubstancia o conhecimento, enquanto capacidade de agir. Por essa razão, o acesso à informação é cada vez mais premente, na medida em que ela constitui a base de apoio para qualquer decisão, em particular numa altura em que se criam formas completamente novas de interacção económica e social num espaço sem fronteiras.

Entre essa informação deve ser destacada a informação estatística, a qual pode ser considerada imprescindível pela sua relevância a nível da tomada de decisões nos mais variados campos. Desta forma, a estatística é reconhecida como sendo um elemento da maior importância na avaliação e na previsão, o que, sendo válido para qualquer país, não o poderia

deixar de ser no âmbito da União Europeia a qual, dispondo de um órgão central de estatística, tem aprovado diversos regulamentos e directivas com o objectivo de fomentar a livre circulação de informação, dentro e entre os Estados-membros.

Se recuarmos um pouco no tempo, e levarmos em linha de conta alguns dos que foram os objectivos presentes ao protocolo estabelecido em 1999 entre o Ministério da Ciência e do Ensino Superior (MCES) e o Instituto Nacional de Estatística (INE) no que diz respeito ao Acesso a Bases de Dados para fins de Investigação Científica, facilmente através dos seus considerandos nos damos conta de dois aspectos:

1º) a informação estatística que em Portugal é produzida pelo INE, sendo em muitos casos a única disponível no que diz respeito à situação económica e social do país, tem fundamental importância no desenvolvimento das actividades de investigação;

2º) esse facto é a razão pela qual o tratamento de dados e de informação estatística em poder do INE é considerado de interesse público, ao ter por objectivo a produção do conhecimento com a finalidade de, com base nele, se tomarem decisões correctas, quer do ponto de vista político, quer económico ou social.

Entre outros, foram estes dois considerandos que estiveram na base do referido protocolo de cooperação entre as duas instituições, e que de ponto de vista prático viria a permitir aos investigadores e para efeitos de investigação científica, o acesso facilitado a bases de dados detidas pelo INE.

Estipula o citado protocolo, que sempre que essas bases contenham informação de natureza individual, a mesma será anonimizada, podendo ter acesso a elas todos os investigadores que desenvolvam a sua actividade de investigação em instituições de investigação científica e desenvolvimento tecnológico credenciadas pelo Ministério da Ciência e do Ensino Superior, tal foi a importância que se reconheceu à necessidade de divulgação da informação para esses fins.

Com o princípio da administração aberta ou do arquivo aberto que se consubstancia no direito de acesso dos cidadãos aos arquivos e registos administrativos (consagrado na Constituição da República Portuguesa de 1976) e mais tarde transposto para Lei através do Código do

Procedimento Administrativo, a informação deixou de ser monopólio do Estado, transformando-se num bem acessível a todos os cidadãos, permitindo dessa forma vir a dar resposta às necessidades de uma sociedade baseada no conhecimento e na informação.

Há no entanto que definir os limites do direito à informação, procurando a todos os níveis encontrar o desejado equilíbrio entre liberdade e responsabilidade. Para tal, a estatística enquanto órgão privilegiado de compilação de informação, tem necessariamente que assentar em normas e/ou mecanismos que impeçam o uso abusivo dessa mesma informação, ao mesmo tempo que asseguram a sua confidencialidade. Assim, em qualquer uma das fases do processo: recolha, tratamento e difusão da informação, há que obedecer a regras que, garantindo a confidencialidade da informação, dêem credibilidade a todo o sistema estatístico.

Surge desta forma o segredo, por um lado como uma excepção à liberdade de informação, e por outro como uma salvaguarda da privacidade dos cidadãos de modo a garantir a confiança no próprio sistema estatístico. Só a garantia de não ver ameaçada a sua privacidade e liberdade, quer enquanto cidadãos, quer enquanto agentes económicos, pode levar no processo de recolha a uma maior adesão dos respondentes, e conseqüentemente a estatísticas fiáveis que preencham os seus verdadeiros objectivos. Aliada à confidencialidade dos dados de carácter individual, a elaboração de estatísticas fiáveis sempre terá por base a confiança no sistema estatístico.

Aparentemente surge aqui uma contradição, na medida em que a salvaguarda da confidencialidade da informação pode *a priori* ser entendida como uma negação da acessibilidade à informação: Contudo, "*Confidencialidade não é sinónimo de negação de acessibilidade*" (CSE, 1999). Há é que ter em conta com o devido cuidado os aspectos acabados de referir, quer no que respeita à recolha dos dados, quer no que concerne ao seu tratamento e respectiva difusão, de modo a garantir que a liberdade de acesso à informação não colida de forma alguma com a vida privada de cidadãos e agentes económicos.

2. O Sistema Estatístico Nacional e o segredo estatístico

Em Portugal, a Lei nº.6/89 de 15 de Abril que aprovou as Bases Gerais do SEN – Sistema Estatístico Nacional (e do qual fazem parte o Conselho Superior de Estatística e o Instituto Nacional de Estatística), consagra:

- (i) no seu artigo 3º. e como funções exclusivas do INE, as de notação, apuramento, coordenação e difusão de dados estatísticos oficiais;
- (ii) no ponto 1 do seu artigo 5º. o princípio do segredo estatístico como princípio fundamental do SEN, tendo em mente salvaguardar a privacidade dos cidadãos, preservar a concorrência entre os agentes económicos, ao mesmo tempo que garante a confiança dos informadores no sistema estatístico nacional.

No entanto, no ponto 5 deste último artigo, a Lei nº.6/89 prevê a possibilidade de, mediante análise caso a caso, poder ser concedida autorização pelo Conselho Superior de Estatística à cedência de dados sobre cooperativas, empresas públicas e privadas, instituições de crédito e demais agentes económicos, desde que estejam em causa necessidades de planeamento e coordenação económica ou relações económicas externas. Ou seja, esta é uma das situações para as quais foi então prevista quebra de sigilo.

Na sua 60ª. Deliberação de 28/04/93, o Conselho Superior de Estatística (CSE) enquanto órgão que superiormente orienta e coordena o Sistema Estatístico Nacional, aprovou o Regulamento de Aplicação do Princípio do Segredo Estatístico, diploma que no seu artigo 21º., regulamenta a libertação do segredo, fazendo então depender o fornecimento da informação abrangida pelo segredo estatístico de autorização expressa do Conselho Superior de Estatística. Este Conselho poderia dar parecer favorável à libertação, desde que os pedidos fossem devidamente fundamentados, e entre outras exigências, expressassem o compromisso das entidades requerentes da libertação do segredo de guardarem absoluto sigilo das informações fornecidas, utilizando-as apenas para os fins mencionados.

Isso significava o condicionar da libertação do segredo a um conjunto de requisitos a que os pedidos teriam de obedecer, e com a particularidade de “só ser permitido divulgar dados que

individualmente se reportassem a, pelo menos três unidades estatísticas de tabulação". Ora, esta regra relativamente ao tipo de informação divulgada (ou cedida), é bastante restritiva do ponto de vista de quem pretende fazer investigação, criando dificuldades óbvias aos utilizadores da informação ao inviabilizar as necessárias desagregações.

Esse Conselho estipula no nº.4 (Capítulo A) da sua 188ª. Deliberação datada de 7 de Julho de 2000, o seguinte: poderão ser considerados favoravelmente os pedidos de libertação do segredo estatístico, devidamente fundamentados e enquadrados nas excepções previstas na Lei, desde que obedeçam a determinadas condições. Entre essas condições, encontra-se especificada a seguinte: *"Tratar-se de informação necessária à prossecução de fins científicos"*, desde que relativa à elaboração de estatísticas comunitárias, nos termos dos artigos 17º e 18º do Regulamento Comunitário (CE) nº. 322/97 de 17 de Fevereiro. De facto, essa Deliberação do CSE, acolhe e transcreve as orientações desse Regulamento Comunitário, incluindo no ponto 4.3, a investigação científica como uma das excepções que permite o levantamento do segredo estatístico.

Note-se no entanto, que anteriormente a essa deliberação do CSE, na síntese efectuada pelo Grupo de Trabalho para Análise e Reflexão sobre as Normas do Instituto do Segredo Estatístico (CSE, 1999), já tinha sido feita referência ao facto da lei não prever a possibilidade de serem utilizados dados nominativos (microdados) para fins científicos ou de investigação. Foi então deixada em aberto, a possibilidade de com maior profundidade serem abordadas pelo Conselho algumas questões, entre as quais a hipótese de virem a ser facultados dados nominativos para efeitos de investigação científica, com a necessária adopção das adequadas regras de controlo de confidencialidade, por forma a salvaguardar o princípio da confiança dos informadores no sistema estatístico. Essa possibilidade teria obviamente de ter em conta o disposto na Lei nº. 67/98 de 26 de Outubro – Lei da Protecção de Dados Pessoais¹.

¹ Esta Lei transpõe para a ordem jurídica portuguesa a Directiva 95/46/CE, do Parlamento Europeu e do Conselho, de 24 de Outubro de 1995, relativa à protecção das pessoas singulares no que diz respeito ao tratamento dos dados pessoais e à livre circulação desses dados.

3. O segredo estatístico face à regulamentação comunitária. Seus efeitos sobre a investigação científica

Conforme vimos anteriormente, a Lei nº.6/89 consagra em Portugal o princípio do segredo estatístico como princípio fundamental do SEN, havendo que assegurar a máxima qualidade, confidencialidade e transparência do sistema, o que passa pela protecção de dados pessoais. Nessa lei, estão definidas as excepções admitidas como sendo susceptíveis de poderem desencadear o processo de levantamento do segredo estatístico.

A nível comunitário, é o Regulamento (CE) 322/97 de 17 de Fevereiro que consagra o acesso a dados confidenciais obtidos para a elaboração de estatísticas comunitárias, desde que estejam em causa finalidades científicas, alargando desta forma o âmbito das excepções admitidas pela lei do SEN.

O Regulamento 322/97/CE, ao estabelecer no seu artigo 10º. os princípios por que devem reger-se as estatísticas comunitárias, nomeadamente princípios de imparcialidade, fiabilidade, pertinência, relação custo/eficácia e transparência, adiciona-lhes o princípio do segredo estatístico. Tendo por objectivo criar um enquadramento normativo para produção das estatísticas comunitárias, este Regulamento no Considerando 13 do seu preâmbulo faz referência ao segredo estatístico da seguinte forma, e passamos a citar: *“Considerando que, a fim de ganhar e manter a confiança dos responsáveis pelo fornecimento dessa informação, é importante proteger os dados confidenciais que as autoridades estatísticas nacionais e comunitárias devem coligir para a produção das estatísticas comunitárias e que a confidencialidade dos dados estatísticos deve manter o mesmo conjunto de princípios em todos os Estados-membros”*.

Começando no seu artigo 13º. por considerar confidenciais dados utilizados pelas autoridades nacionais e comunitária para a produção de estatísticas comunitárias, desde que os mesmos permitam a identificação directa ou indirecta das unidades estatísticas, razão pela qual revelarão informações de carácter individual, o artigo 17º. do Regulamento 322/97/CE faz referência ao facto de se poder aceder com fins científicos a dados confidenciais, o que pode ser concedido pela autoridade nacional responsável pela sua produção (no nosso caso, será obviamente o INE) desde que o nível da sua protecção esteja garantido por normas que o

mesmo Regulamento estabelece no seu artigo 18º., e que visam assegurar a protecção dos dados de natureza confidencial e evitar riscos de divulgação ilícita.

Parece-nos ser de realçar o facto de, apesar da alusão aos fins científicos como justificando o levantamento do segredo, não ser feita em todo o Regulamento 322/97/CE, qualquer referência explícita a dados nominativos.

Esta questão dos dados nominativos parece-nos de sobremaneira importante, na medida em que as normas relativas à cedência de dados confidenciais, até há bem pouco tempo, por não contemplarem microdados, revelavam-se bastante restritivas para quem faz investigação.

O Regulamento (CE) nº. 831/2002 de 17 de Maio que implementa o Regulamento 322/97/CE, sendo relativo às estatísticas comunitárias no que respeita ao acesso a dados confidenciais para fins científicos, estabelece as normas de acesso a dados confidenciais transmitidos à autoridade comunitária e as regras de cooperação entre essa autoridade e as nacionais, de forma a facilitar esse acesso. E vem quanto a nós, permitir viabilizar alguma da investigação que era vedada aos investigadores por falta de desagregação da informação anteriormente disponibilizada. Vejamos como.

No seu Artigo 2º. esse Regulamento define:

- (i) «o acesso a dados confidenciais» como sendo *“o acesso nas instalações da autoridade comunitária a microdados tornados anónimos ou a sua divulgação”*,
- (ii) e «microdados tornados anónimos» como sendo *“registos estatísticos individuais que foram modificados de modo a minimizar, de acordo com a actual melhor prática, o risco de identificação das unidades estatísticas a que se referem”*.

Determina ainda no seu Artigo 3º., que o acesso a dados confidenciais pode ser concedido pela autoridade comunitária a investigadores, desde que pertencentes a Universidades e Organizações ou Instituições de Investigação Científica, mediante pedidos devidamente fundamentados com apresentação das respectivas propostas de investigação.

Quanto a nós, este Regulamento vem de encontro àqueles que são os interesses da investigação científica, ao estabelecer:

- quais as entidades que podem ter acesso a dados confidenciais para fins científicos,
- qual a forma desse acesso,
- com a possibilidade de acederem a conjuntos de microdados, desde que previamente se proceda à sua anonimização.

Esta questão da anonimização dos dados parece-nos de facto um aspecto importante a realçar, pelas razões que passamos a explicar.

Na caracterização de uma dada população, a estatística é elaborada a partir de recolha de informação respeitante aos indivíduos que a constituem, pelo que podemos afirmar serem eles, os indivíduos, a base de toda a informação. Apesar disso, os resultados estatísticos nada revelam sobre as unidades individuais fonte de informação, razão pela qual, ao condensarem anonimamente dados individuais, nunca é colocado nestes casos qualquer problema em matéria de protecção de dados. Significa isto, que a atenção da estatística está canalizada para a informação quantitativa respeitante a grupos ou conjuntos, que muito embora seja resultante de observações individuais, revela um interesse que é independente dos atributos daquelas que foram individualmente as unidades de observação.

Do ponto de vista da investigação científica, o que se passa é o seguinte: existem situações de análise que pela sua natureza podem exigir a utilização de microdados, porém sem qualquer necessidade, e muito menos qualquer interesse por parte do investigador em identificar as unidades inquiridas, razão pela qual, o recurso ao anonimato serve os objectivos dos investigadores, garantindo em simultâneo que não seja posta em causa a privacidade dos titulares dos dados.

4. Relato de uma experiência. Considerações finais

Antes de finalizarmos esta apresentação, fazemos questão de efectuar um breve relato de uma experiência partilhada com o INE e o CSE, no que respeita a dois pedidos de informação estatística que tiveram lugar no âmbito de dois projectos de investigação científica. Deve salientar-se que os projectos então apresentados em reunião da Secção Permanente do

Segredo Estatístico, obtiveram parecer favorável de todos os presentes no que respeita à cedência da informação solicitada, desde que anonimizada.

Relativamente àquilo que no decorrer do processo, foi considerado por nós investigadores como alguma burocracia e falta de celeridade, podemos agora dizer honestamente serem de mais fácil entendimento, talvez por estarmos por dentro de todos os requisitos e formalidades a cumprir em caso de pedidos de dados abrangidos pelo segredo estatístico, por forma a estarem preenchidas as necessárias condições que garantam a confidencialidade dos dados e a confiança no sistema estatístico.

Porém, se a demora na cedência da informação se entende pelos motivos aludidos, existe um aspecto que fazemos questão de referir, muito embora não tenhamos conhecimento se o mesmo pode ser torneado, o que permitiria eventualmente tornar mais célere a resposta a pedidos futuros. O facto de, quando da primeira cedência de dados se terem verificado alguns problemas de incompatibilidade em termos do *software* do INE em Lisboa, da Direcção Regional do Alentejo e da Universidade de Évora, levou a que alguns ficheiros tivessem de ter sofrido partição e algum tratamento bastante moroso. Isso constituiu um problema para os investigadores, e muito embora não tenha directamente a ver com a questão da libertação da informação, esse facto forçou-nos a apresentar um segundo pedido de actualização ao INE, dada a demora na leitura da informação disponibilizada. E quanto a nós, problemática foi a repetição de todo o processo, numa situação em que aquilo que se pretendia era uma actualização dos dados pura e simples, na continuação de um dos dois projectos de investigação, já anteriormente apresentado em reunião da Secção Permanente do Segredo Estatístico.

Face a isto, a questão que nos parece pertinente é a seguinte: tratando-se do mesmo projecto, e da mesma informação, apenas para um ou dois anos posteriores, a repetição do processo na sua totalidade não poderia ser evitada? Para quem faz investigação, e em particular quando o tempo que para ela se disponibiliza tem de ser compatibilizado com calendários lectivos, a demora de dois ou três meses pode fazer imensa diferença. Desconhecendo se este é um problema que pode ou não ser ultrapassado, não quisemos deixar de o referir, na expectativa de para ele poder ser encontrada uma solução.

Um aspecto final que não queremos deixar de frisar, diz respeito ao facto de nunca ter havido da parte dos investigadores envolvidos, ao solicitarem dados por empresa, e como tal, ao abrigo do segredo estatístico, qualquer intuito (porque completamente desnecessário) de identificar as unidades inquiridas. Logo, o recurso ao anonimato sempre serviu os nossos objectivos, como serve decerto os de todos os outros investigadores, ao mesmo tempo que permite garantir a privacidade dos titulares dos dados. O que no nosso caso tornou um pouco mais complicada a anonimização, foi o facto de ter sido solicitada a libertação do segredo relativamente a dados obtidos por amostragem em dois inquéritos diferentes, nomeadamente o Inquérito às Empresas Harmonizado (IEH) e o Inquérito Anual à Produção Industrial (IAPI), com o objectivo de efectuar cruzamento entre ambos. Este facto implicou o estabelecimento de correspondência entre os dois inquéritos, forçando o INE ao anonimizar os dados, a ter de efectuar o cruzamento da respectiva informação. No entanto, esse cruzamento não constituiu obstáculo à cedência da informação com total garantia da sua anonimização.

Por tudo o que foi exposto, entendemos que com a legislação em vigor, o segredo estatístico não deve mais ser entendido como condicionante da investigação científica, por estarem criadas as condições do ponto de vista legal para que possa ser fornecida informação confidencial aos investigadores, até mesmo naqueles casos que eram à partida excluídos por envolvimento de dados nominativos. Com isso, a investigação científica só terá a ganhar. Em particular no que diz respeito à investigação científica portuguesa, quem sabe, talvez esse facto possa vir a contribuir para o seu crescimento, e conseqüentemente para a redução do fosso que, a despeito da evolução verificada nos últimos anos, ainda existe entre a produção científica nacional e aquela que tem lugar na maioria dos países da União Europeia.

Referências

Conselho Superior de Estatística (CSE), 60^a. Deliberação de 28 de Abril de 1993.

Conselho Superior de Estatística (CSE), *“1.º Relatório do Grupo de Trabalho para Análise e reflexão sobre as Normas Actuais do Instituto do Segredo Estatístico”*, DOCT/317/CSE/SE – 2, Lisboa, 24 de Maio de 1999.

Conselho Superior de Estatística, 188^a. Deliberação de 7 de Julho de 2000.

Diário da República, I Série, Lei nº.6/89 de 15 de Abril - Lei do Sistema Estatístico Nacional.

Diário da República, I Série, Lei nº. 67/98 de 26 de Outubro – Lei da Protecção de Dados Pessoais.

Gabinete de Estudos e Planeamento do Ministério da Justiça, *“Regulamento de Aplicação do Princípio do Segredo Estatístico”*.

Jornal Oficial das Comunidades Europeias, nº L 052, Regulamento (CE) nº. 322/97 de 17 de Fevereiro.

Jornal Oficial das Comunidades Europeias, nº L 133, Regulamento (CE) nº. 831/2002 de 17 de Maio.

Ministério da Ciência e do Ensino Superior e Instituto Nacional de Estatística, *“Protocolo para Acesso a Bases de Dados para fins de Investigação Científica”*, Lisboa, Outubro de 1999.