

CATÓLICA LAW REVIEW

VOLUME I
n.º 2
mai. 2017

DIREITO PRIVADO

Carlos Ferreira de Almeida

Joseph William Singer

Paulo Mota Pinto

Henrique Sousa Antunes

Júlio Gomes



CATÓLICA

RESEARCH CENTRE
FOR THE FUTURE OF LAW

LISBOA - PORTO

CATÓLICA LAW REVIEW

VOLUME I \ n.º 2 \ maio 2017

DOCTRINA

Carlos Ferreira de Almeida

Professor Catedrático Jubilado
Faculdade de Direito da Universidade Nova de Lisboa

Joseph William Singer

Bussey Professor of Law
Harvard Law School

Paulo Mota Pinto

Professor Auxiliar
Faculdade de Direito da Universidade de Coimbra

COMENTÁRIO DE JURISPRUDÊNCIA

Henrique Sousa Antunes

Professor of Law
Católica Lisbon Law School, Universidade Católica Portuguesa

RECENSÃO

Júlio Gomes

Juiz Conselheiro (STJ)
Professor Convidado da Escola de Direito no Porto
da Universidade Católica Portuguesa

UNIVERSIDADE CATÓLICA EDITORA

Proprietário

Católica Research Centre for the Future of Law
Católica Law Review
VOLUME I \ n.º 2 \ maio 2017

Conselho de Direção

Rui Medeiros
Catarina Santos Botelho
Elsa Vaz de Sequeira
Gonçalo Almeida Ribeiro
Maria d'Oliveira Martins
Sofia Pais

Redação e Administração

Católica Research Centre for the Future of Law
Faculdade de Direito
Universidade Católica Portuguesa
Palma de Cima 1649-023 Lisboa
Tel. (351) 21 721 41 78 | Fax. (351) 21 721 41 77
catolicallawreview.fd@ucp.pt | www.fd.lisboa.ucp.pt

Este trabalho é financiado por fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto Ref.ª UID/DIR/04859/2013

© Universidade Católica Editora

Revisão Editorial	António Brás
Capa e conceção gráfica	Ana Luísa Bolsa 4 ELEMENTOS
Paginação	acentográfico
Impressão	Sersilito – Empresa Gráfica, Lda.
Depósito Legal	420225/17
Data	maio 2017
Tiragem	300 exemplares
Periodicidade	quadrimestral
ISSN	2183-9336



Universidade Católica Editora
Palma de Cima 1649-023 Lisboa
Tel. (351) 217 214 020 | Fax. (351) 217 214 029
uce@uceditora.ucp.pt | www.uceditora.ucp.pt

ÍNDICE

7 Nota da Direção

DOCTRINA

9 Invalidade, inexistência e ineficácia

Carlos Ferreira de Almeida

Resumo: Este texto trata de três conceitos – invalidade, inexistência e ineficácia de atos jurídicos – e das relações entre eles. O estudo não tem um desiderato prático, mas como a lei é construída com palavras, os regimes muitas vezes dependem dos conceitos usados pelas regras legais. Uma discussão relevante derivada dos conceitos e das regras é que, em sentido estrito, o chamado “contrato inválido” não é um contrato.

Palavras-chave: Atos legislativos. Invalidade. “Contrato inválido”. Inexistência. Ineficácia *stricto sensu*.

Abstract: This essay concerns three concepts – invalidity, non-existence and inefficacy of legal acts – and the relations between them. The study has no practical goals but, as law is constructed with words, the regimes often depend on the concepts used by the legal rules. One relevant conclusion derived from concepts and from rules as well is that, strictly speaking, the so called “invalid contract” is not a contract.

Keywords: Legal acts. Invalidity. “Invalid contract”. Non-existence. Inefficacy *stricto sensu*.

35 We don't serve your kind here: public Accommodations and the mark of Sodom

Joseph William Singer

Abstract: As we celebrate the fiftieth anniversary of the Civil Rights Act of 1964, we ask: can one person – because of his sexual orientation, race, gender or religious beliefs – be disabled from accessing services that human beings need to live, including food, shelter, bathrooms and fuel?

Even though such discrimination is not socially acceptable by today's standards, it is wrong to assume that there is no need for laws that prohibit discriminatory conduct in this day and age. Today the question is not whether one can find a store willing to let you in and treat you with dignity.

The question is whether one has a right to enter stores without worrying about such things.

Keywords: Civil rights. Limits of private property. Public accommodation laws. Equal access to goods in the marketplace. Segregation. Discrimination. Right to select people to do business with.

Resumo: No momento em que se celebra o décimo quinto aniversário do Civil Rights Act de 1964, perguntamo-nos: pode alguém ser discriminado – por causa da sua orientação sexual, raça, género ou convicções religiosas – no acesso a serviços que todos os seres humanos precisam para viver, como alimento, abrigo, instalações sanitárias ou combustível?

Mesmo que hoje tal discriminação não se coadune com os parâmetros de uma sociedade moderna, é errado assumir que não há necessidade de leis que proíbam os comportamentos discriminatórios nos tempos e dias que correm. Hoje, a questão não é a de saber se se pode encontrar um estabelecimento disponível para o deixar entrar e tratá-lo com dignidade. A questão é de saber se se tem o direito de entrar nos estabelecimentos sem se preocupar com essas coisas.

Palavras-chave: Direitos civis. Limites ao direito de propriedade. Leis de estabelecimentos abertos ao público. Acesso igual a bens no mercado. Segregação. Discriminação. Direito a selecionar as pessoas com quem contratar.

- 59 Sobre a alegada “Superação” do Direito pela análise económica (ilustrada com a análise das medidas da indemnização contratual)

Paulo Mota Pinto

Resumo: Os problemas fundamentais do direito dos contratos são problemas especificamente jurídico-normativos, devendo por isso ser tratados, não só de acordo com as referências próprias de cada sistema, mas também de acordo com a intenção de sentido que inspira a perspectiva jurídica.

Os conhecimentos científicos e a técnica da análise económica do direito devem ter um papel auxiliar para tornar clara a relevância da realidade a que se dirigem, a estrutura factual, tanto desta, como dos efeitos de um determinado regime. A não ser assim estaremos perante um sério desvio metodológico, a revelar grave incompreensão do específico sentido do Direito, quando não mesmo uma opção antropológica discutível e, se levada até ao fim, um retrocesso cultural.

A rejeição do instrumentalismo da análise económica do Direito tem reflexos no modo como se encara a própria decisão do devedor cumprir ou não cumprir obrigações. A esta luz, aborda-se criticamente a teoria do não cumprimento eficiente.

Palavras-chave: Análise económica do direito. Direito Privado. Perda de autonomia do direito. Eficiência. Cumprimento e não cumprimento das Obrigações. Responsabilidade contractual. Não cumprimento eficiente.

Abstract: The fundamental problems of contract law are specifically legal and normative, and should therefore be dealt with, not only according to the specific references of each system, but also in accordance with the meaning intended by the legal perspective.

Scientific knowledge and the technique of the law & economics must have an auxiliary role to make clear the relevance of the reality to which they are addressed, the factual structure and the effects of a given regime. If not so, we will be facing a serious methodological deviation, revealing a serious misunderstanding of the specific sense of the law. If it is not a debatable anthropological option it is, if carried to the end, a cultural regression.

The rejection of the instrumentalism of the law & economics has consequences in the way the debtor will comply or not his obligations. As such, the theory of efficient breach is critically addressed.

Keywords: Law & economics. Private law. Law's loss of autonomy. Efficiency. Compliance and non-compliance with legal obligations. Contractual liability. Efficient breach.

COMENTÁRIO DE JURISPRUDÊNCIA

- 95 Intellectual property rights: infringement and calculation of damages
Henrique Sousa Antunes

RECENSÃO

- 111 *Die Richtigkeitsgewähr des Tarifvertrags*
Júlio Gomes

Nota da Direção

O segundo número da *Católica Law Review* apresenta-se com cinco textos de elevadíssima qualidade científica. Inserindo-se ainda no ano inicial da revista, optou-se por, à semelhança do que sucedeu com o primeiro número, não fazer uso do recurso a *call for papers*. Todos os artigos são, pois, fruto de convites dirigidos aos seus autores.

O presente número conta, na parte doutrinal, com os contributos de Carlos Ferreira de Almeida (Nova-Lisboa), Joseph Singer (Harvard Law School) e Paulo Mota Pinto (Coimbra). Inclui ainda um comentário a uma decisão do Tribunal de Justiça da União Europeia feita por Henrique de Sousa Antunes (Católica-Lisboa) e uma revisão de Júlio Gomes (Católica-Porto).

A revista reúne assim artigos de Professores de reconhecido mérito, agregando temas variados e de grande atualidade. O texto de Carlos Ferreira de Almeida é dedicado à análise dos conceitos de invalidade, inexistência e ineficácia dos atos jurídicos e da relação que se estabelece entre eles. A reflexão de Joseph Singer, tratando dos limites ao direito de propriedade, aborda o problema da discriminação, questionando a existência de um direito a selecionar as pessoas com quem contratar. O estudo de Paulo Mota Pinto propõe uma visão crítica da técnica da análise económica do direito, com consequências práticas no âmbito da compreensão do cumprimento e não cumprimento das obrigações.

Sendo um número de direito privado, é importante notar que qualquer um dos textos que nele se contém aborda temas transversais, e que, por isso, facilmente podem cativar a atenção de juristas mesmo fora dessa área.

Editorial Note

The second issue of *Católica Law Review* brings us five articles of very high scientific quality. In this first year of the *Católica Law Review*, it was the Editors' decision not to call for papers. Therefore, all articles are the result of invitations addressed to their authors.

In the doctrinal part, this issue includes texts of Carlos Ferreira de Almeida (Nova-Lisboa), Joseph Singer (Harvard Law School) and Paulo Mota Pinto (Coimbra). Henrique de Sousa Antunes (Católica-Lisbon) provides a comment on a decision of the European Court of Justice and Júlio Gomes (Católica-Porto) presents us a book review.

The *Católica Law Review* gathers articles of well-known teachers about current wide-ranging themes. In his article, Carlos Ferreira de Almeida analyses the concepts of invalidity, inexistence and inefficacy of legal acts and the relationship established between themselves. Joseph Singer's reflection deals with the limits to property rights, addressing the problem of discrimination and questioning the existence of a right to select the people to enter business with. Paulo Mota Pinto proposes a critical view of the law & economics approach, with practical consequences in the understanding of legal obligations.

Even though this issue is dedicated to private law, it is important to note that all articles address transversal matters, and that therefore they should be under the attention of lawyers from different areas.

Invalidade, inexistência e ineficácia

Carlos Ferreira de Almeida

Professor Catedrático Jubilado
Faculdade de Direito da Universidade Nova de Lisboa

SUMÁRIO

1. Âmbito do estudo
2. Ineficácia e invalidade
3. Inexistência
4. Caracterização do "contrato nulo"
5. Ineficácia em sentido restrito
6. Referentes da invalidade, da inexistência e da ineficácia *stricto sensu*

1. Âmbito do estudo

A ineficácia, a invalidade e a inexistência são questões da teoria do direito e, portanto, comuns a todos os ramos de direito. Neste estudo, o *corpus* é mais limitado, porque tem por referência primordial contratos regidos pelo direito privado, o que implica, por força da técnica do Código Civil, que, nalguns pontos, o *corpus* se alargue aos negócios jurídicos¹. Quando seja possível e seguro, usarei a generalização para qualquer ato jurídico.

Ficam assim, em regra, fora desta análise referências específicas a regras de direito público, designadamente as que são aplicáveis aos tratados internacionais², às leis³, aos regulamentos e aos atos administrativos⁴, aos contratos públicos⁵, aos registos públicos⁶, às nulidades em processo penal⁷ e em processo civil⁸.

1 Na verdade, uma leitura atenta e descomprometida da parte geral do Código Civil conduz à descoberta frequente (para não dizer normal) do contrato escondido por detrás do negócio jurídico e da declaração negocial: cfr. FERREIRA DE ALMEIDA (2013a), pp. 16 ss, 40 s.

2 A invalidade, a cessação da vigência e a suspensão dos tratados são reguladas pelos artigos 42.º a 72.º da Convenção de Viena de 1969 sobre o Direito dos Tratados, em vigor desde 1980 (em Portugal desde 2003); a ineficácia inicial dos tratados verifica-se no período compreendido entre a adoção do texto (artigo 9.º) e a sua entrada em vigor (artigo 24.º), designadamente por ratificações ou adesões insuficientes; as reservas (artigos 19.º a 23.º) determinam ineficácia parcial. Ver, por todos, DÖRR & SCHMALENBACH (2012).

3 As leis, além de inexistentes (cfr. nota 38 *in fine*), podem sofrer de invalidade por inconstitucionalidade – nulidade, para GOMES CANOTILHO (2003), pp. 947 ss, p. 956; anulabilidade *ex tunc*, para MEDEIROS (1999), pp. 538 ss, 551 ss, 651 ss – ou de ineficácia, designadamente por falta de publicação (GOMES CANOTILHO, 2003), p. 842) ou se a lei for temporária ou se ficar suspensa por efeito de tratado internacional que trate da mesma matéria: BRITO (1999), pp. 438 ss, citando GALVÃO TELES, M. (1967), pp. 98 ss.

4 Quanto à invalidade, Código do Procedimento Administrativo (aprovado pelo Decreto-Lei n.º 4/2015, de 7 de janeiro), artigos 143.º e 144.º, 161.º a 164.º O mesmo Código regula alguns aspetos da ineficácia dos regulamentos (artigos 145.º e 146.º sobre caducidade e revogação) e da ineficácia dos atos administrativos (cfr. artigos 165.º e seguintes sobre revogação; artigo 149.º sobre condição e termo; artigo 157.º sobre eficácia diferida ou condicionada; artigos 139.º e 158.º, n.º 2, sobre a publicação e sua falta; artigo 160.º sobre oponibilidade a partir da notificação).

5 A invalidade e a (in)eficácia dos contratos públicos são reguladas pelo respetivo Código nos artigos 283.º a 285.º e 287.º; ver PACHECO DE AMORIM (2008), pp. 627 ss.

6 Sobre a inexistência, a nulidade e o cancelamento dos registos públicos, ver Código do Registo Civil, artigos 85.º a 91.º; Código do Registo Predial, artigos 14.º a 17.º, 56.º a 59.º, 101.º, n.º 1, alínea f). O Código da Propriedade Industrial regula a nulidade e a anulabilidade das patentes, modelos de utilidade e registos (artigos 33.º a 36.º e *passim*), a sua caducidade (artigos 37.º e *passim*), a renúncia (artigo 38.º), o esgotamento (artigos 103.º, 146.º, 166.º, 205.º, 259.º) e a inoponibilidade de direitos (artigos 104.º, 147.º, 167.º). Quanto à invalidade do registo (privado) de valores mobiliários escriturais, cfr. FERREIRA DE ALMEIDA (2005a), pp. 940 ss.

7 O respetivo Código inclui um regime geral de nulidades (artigos 118.º a 122.º) e comina com nulidade vários atos; a nulidade por ilicitude da obtenção da prova – COSTA ANDRADE (1991) – assume dignidade constitucional (artigo 32.º, n.º 8).

8 A nulidade dos atos processuais e os vícios da sentença, incluindo a nulidade, são regulados pelos artigos 186.º a 202.º e 613.º a 615.º do respetivo Código; cfr. LEBRE DE FREITAS (2008), pp. 342 ss, e LEBRE DE

Espero, porém, que a visão civilista possa dar algum contributo para o estudo destes regimes. Para as invalidades, o direito civil pode ser supletivo ou, pelo menos, inspirador – por coincidência, semelhança, especialidade ou até por contraste – das instituições de direito público. Em relação à inexistência e à ineficácia, desprovidas de regime civil comum, os fragmentos recolhidos no Código Civil e noutras fontes dispersas de direito privado podem também aproveitar ao direito público.

Este estudo não tem preocupações pragmáticas diretas. Mas na ciência jurídica, como em qualquer ciência, conceitos claros e coerentes não devem ser desprezados só porque se não vislumbra uma imediata aplicação prática. Além da relevância taxinómica por si só, resta sempre a possibilidade de algum chamamento inesperado.

O tema deste artigo surge em todos os manuais de teoria geral do direito civil. Mas, sob uma aparente convergência, não são afinal poucas as questões controvertidas ou incompletamente exploradas. Dar-me-ei por satisfeito se resolver algumas e se suscitar novas dúvidas.

2. Ineficácia e invalidez⁹

I. *Ineficácia*. Um ato é ineficaz sempre que não produza todos ou parte dos efeitos que a categoria a que pertence está, em abstrato, apta para produzir¹⁰.

Sendo performativo, como é o negócio jurídico¹¹, a sua essência – efeitos conformes ao significado – é afetada se tal conformidade não se verificar. A desconformidade resulta de algum elemento perturbador, intrínseco ou extrínseco, impeditivo da aplicação das regras constitutivas da eficácia. A ineficácia jurídica é o equivalente do insucesso (*infelicity*, *unhappiness*) na filosofia analítica¹².

Ineficácia não equivale necessariamente à falta total de efeitos. No seu sentido mais amplo, a ineficácia de um ato jurídico verifica-se sempre que os efeitos

FREITAS (2001), pp. 663 ss. Sobre a nulidade por ilicitude da prova, ver ALEXANDRE (1998) e SALAZAR CASANOVA (2004).

9 Sobre a evolução dos conceitos, MENEZES CORDEIRO (2008), pp. 233 ss, aprox. = MENEZES CORDEIRO (2014), pp. 779 ss.

10 Outras formulações com pontos coincidentes ou próximos: se faltarem as consequências que normalmente deveria produzir ou que eram razoavelmente esperadas, correspondentes ao tipo normal do negócio jurídico – DIEZ-PICAZO (2007), pp. 555, 571; se o contrato não produzir os efeitos que estava tendencialmente destinado a produzir – LASARTE (2006), p. 166; se não tiver os efeitos jurídicos pretendidos – WOLF & NEUNER (2012), pp. 666.

11 Cfr. FERREIRA DE ALMEIDA (2013a), pp. 30 ss.

12 Cfr. AUSTIN (1962), pp. 14 e *passim*; SEARLE (1969), p. 54.

próprios do ato não se verifiquem¹³ no todo ou em parte, que não se verifiquem logo ou que já não se verifiquem. A ineficácia é compatível com a produção de outros efeitos derivados do próprio ato ou até com efeitos derivados da ineficácia do ato.

Por exemplo, um contrato de compra e venda anulado por erro é ineficaz na medida em que não gera as obrigações de entregar e de pagar que são típicas daquele contrato. Mas pode ter como efeito a restituição do que tiver sido prestado e a responsabilidade civil pré-contratual da contraparte do errante. Um ato com condição suspensiva não engendra, ou não engendra logo, os efeitos que pode afinal gerar, mas cria imediatamente expectativas tuteladas.

Na perspectiva que aqui se assume, a ineficácia *lato sensu* desdobra-se em três grandes categorias: invalidade, inexistência e ineficácia *stricto sensu*.

II. *Invalidade*. Passado o limiar da existência (ver *infra* n.º 3), o ato que, em abstrato, tenha aptidão para produzir efeitos jurídicos é tendencialmente inválido se violar valores consagrados numa norma legal imperativa. Mas a ilicitude não é critério único de invalidade, porque, por exemplo, são inválidos atos sem a forma bastante e os requisitos de forma são alheios a critérios de licitude. A cominação da invalidade não tem de ser expressa na lei, podendo resultar da sua interpretação com todos os elementos relevantes, incluindo a inserção no sistema jurídico¹⁴.

A invalidade nem sempre tem a natureza de sanção¹⁵ para um comportamento. Tal sucede, sim, se o comportamento for ilícito (ilegalidade material, violação da ordem pública, simulação), mas não se verifica quando a invalidade seja originada pela proteção em função da capacidade jurídica (menor, interdito, pessoa jurídica) ou da plena liberdade de decisão (incapacidade acidental, erro/obstáculo). Nalgumas destas situações, a invalidade afeta também quem tenha agido em plena conformidade com o direito.

A invalidade é uma das fontes da ineficácia, por ser fator impeditivo da conformidade dos efeitos do ato com o seu significado. Invalidade significa não valer, não ter força jurídica, por desconformidade com um valor jurídico¹⁶.

A invalidade é em princípio originária¹⁷ ou genética, no sentido de que se avalia com referência ao tempo da formação do ato, sem prejuízo da sua (nor-

13 GALGANO (2007), p. 351, escreve que o contrato inválido é também ineficaz, porque improdutivo do efeito próprio do contrato.

14 Assim também DIEZ-PICAZO (2007), pp. 559 s, sustentando que a invalidade pode ser reclamada pelos princípios gerais de direito.

15 Como defende DIEZ-PICAZO (2007), *loc. cit.*

16 Note-se a comum raiz etimológica de valer e de valor.

17 OLIVEIRA ASCENSÃO (2003), p. 374.

mal) detecção posterior. A lei aplicável à invalidez é aquela que vigorar ao tempo da prática do ato, mas, se esta lei não for transitória, uma lei posterior que agilize a validade (por exemplo, em relação a requisitos formais, à incapacidade ou à ilicitude), deve aplicar-se para convalidação de atos anteriores, em benefício do princípio do *favor negotii* e da presunção da evolução positiva das políticas legislativas.

Se a invalidez se reportar à ordem pública, aos bons costumes ou a outros conceitos indeterminados mutáveis com o decurso do tempo, a evolução de que resulte a validade deve, por razões semelhantes, prevalecer sobre a invalidez apreciada com os critérios vigentes ao tempo da formação do ato.

Outras desconformidades ou anomalias, suscetíveis de avaliação reportada ao tempo presente, como a impossibilidade superveniente, a alteração das circunstâncias e o incumprimento, não se reportam à validade, mas à relação entre o efeito jurídico e a sua execução prática.

III. *Nulidade e anulabilidade*. A invalidez desdobra-se em duas modalidades: a nulidade e a anulabilidade¹⁸.

O ato nulo¹⁹ não pode valer; o ato anulável é um ato em princípio válido que pode não valer. A nulidade de um ato declara-se, reconhece-se; o ato anulável pode ser anulado ou não, conforme seja ou não exercido o direito potestativo de anulação. O ato anulável é um ato originariamente válido com a validade resolúvel²⁰, embora se possa também dizer que é um ato inválido com a invalidez suspensa²¹. Em qualquer caso, a invalidez, como outras ineficácias, não exclui que do ato decorram efeitos parciais ou diferentes dos que lhes seriam próprios.

Os regimes gerais da nulidade e da anulabilidade diferem quanto aos pressupostos da sua invocabilidade, verificação e pronúncia (artigos 286.º e 287.º²²), mas convergem nas consequências da declaração de nulidade e da anulação

18 No direito francês, usa-se a contraposição equivalente entre nulidade absoluta e nulidade relativa – por todos, FLOUR, AUBERT & SAVALUX (2000), pp. 230 ss – que também era a preferida na vigência do Código de Seabra; cfr. ANDRADE (1964), pp. 415 ss; GALVÃO TELLES, I. (1962), pp. 329 s, embora este já adotasse a nomenclatura atual. Para o direito português atual, uma parte da doutrina – cfr. CARVALHO FERNANDES (2010), pp. 495 ss – distingue entre invalidez absoluta e relativa, correspondentes à nulidade e à anulabilidade, respetivamente.

19 Do latim *nullus*, que significava *nada, ninguém*.

20 No Projeto de Quadro Comum de Referência – *Principles. Definitions and Model Rules of European Contract Law. Draft Common Frame of Reference* (2009), pp. 546, 569 – define-se como nulo (*void*) o ato que automaticamente não tem efeito desde o início e como anulação (*avoidance*) o processo que determina a ineficácia retroativa de um ato que era válido desde o início (ver também a definição de *voidable*).

21 Para DIAS MARQUES (1992), pp. 97 s, a anulabilidade fica a meio caminho entre a invalidez e a validade. O ato anulável pode ser encarado tanto do ponto de vista da validade (precária e resolúvel) como da invalidez (eventual e suspensa).

22 Os preceitos legais a seguir citados sem outra indicação pertencem ao Código Civil vigente.

(artigos 289.º a 293.º). Tanto o ato nulo como o ato anulado podem ser validados por redução e conversão (artigos 292.º e 293.º), mas o ato anulável é, além disso, suscetível de validação por confirmação pelo titular do direito de anulação (artigo 288.º).

Como a legitimidade e o tempo para a declaração de nulidade são mais flexíveis do que a legitimidade e o tempo para a anulação, é frequente dizer-se que o critério da política legislativa para cominar nulidade ou anulabilidade deriva da maior ou menor gravidade do vício ou da violação do interesse público ou geral e de interesse privado ou particular. Mas, se assim for, não é fácil compreender que o vício de forma produza nulidade, mesmo que não afete terceiros, enquanto os atos viciados por dolo ou coação moral são apenas anuláveis. Como o critério legal dos interesses públicos e privados nem sempre se confirma, parece melhor renunciar à sua utilização.

Por isso, e porque não se vislumbra um outro critério material de distinção que se baseie numa destrição clara segundo a natureza ou a categoria dos vícios geradores de invalidade, apenas se pode concluir modestamente que, na alternativa entre as duas espécies de invalidade, a nulidade é o regime regra²³ (cfr. artigo 294.º), aplicando-se a anulabilidade ao conjunto de casos que a lei comina com tal consequência²⁴.

Além do regime comum das invalidades, há regimes especiais de invalidades (cfr. artigo 285.º, 1.ª parte), constantes do Código Civil e de outras leis, que se têm qualificado como atípicas, mistas ou híbridas.

Verdadeiramente atípica é a nulidade dos contratos de jogo e de aposta, que admite uma categoria de contratos nulos, mas lícitos, que são fonte de obrigações naturais (artigo 1245.º).

As restantes invalidades atípicas consistem quase todas em nulidades com regras especiais para a legitimidade ou o tempo da ação de declaração de nulidade, próximas do regime da anulabilidade (por isso, se dizem mistas). Assim, quanto à legitimidade: só o promitente-comprador tem legitimidade para arguir a nulidade por vício de forma ou por omissão da referência à licença de construção ou de utilização em relação ao contrato-promessa a que se refere o artigo 410.º, n.º 3; a nulidade da cessão de créditos litigiosos só pode ser arguida pelo cessionário (artigo 580.º, n.º 2); a nulidade do contrato de edição por vício de forma só pode ser invocada pelo autor (Código do Direito de Autor, artigo 87.º, n.º 2). Quanto à oportunidade: a nulidade do contrato de sociedade comercial só pode ser arguida no prazo de três anos após o registo, com um limitado conjunto de fundamentos (Código das Sociedades Comerciais, artigos 42.º e 44.º).

23 GALVÃO TELLES, I. (1962), p. 366 s.

24 GALGANO (2007), p. 275.

As anulabilidades atípicas são mais circunscritas, mas podem exemplificar-se com o artigo 1640.º, que estende a legitimidade para a requerer a anulação do casamento a um círculo de pessoas que excede o previsto no artigo 287.º, n.º 1. O artigo 13.º, n.º 5, 2.ª parte, do Decreto-Lei n.º 133/2009, de 2 de junho, sobre crédito ao consumo, restringe ao consumidor a invocação da invalidade do contrato por um conjunto de fundamentos, formais e outros, sem distinguir entre nulidade e anulabilidade.

3. Inexistência

I. *Situações de inexistência*²⁵. A ideia de inexistência do contrato surgiu no âmbito do casamento logo após o Código Civil de Napoleão (embora com precedente nos canonistas) como fundamento para ultrapassar a rigidez da regra *pas de nullité sans texte* e, deste modo, recusar qualquer eficácia a situações não previstas na lei como inválidas, designadamente a falta de consentimento dos nubentes, o casamento entre pessoas do mesmo sexo e a ausência de oficial do estado civil.

No atual direito civil francês dos contratos, a inexistência está em declínio na doutrina por se julgar inútil, complicada ou inexata, mas continua a ser usada por alguma jurisprudência e vem sendo reconhecida em direito administrativo e processual²⁶. No direito espanhol, as opiniões são maioritariamente adversas, por inutilidade da ideia perante a de nulidade²⁷. Nos direitos português²⁸ (e italiano²⁹) a categoria é controversa mas aceite pela opinião dominante e profusamente usada pela jurisprudência.

Tendo com referência o direito português vigente, parece-me que o ato jurídico, o negócio jurídico e o contrato são inexistentes num conjunto não homogêneo de situações, que se podem reconduzir às seguintes³⁰:

1.ª Se o facto referido não ocorreu. É a chamada inexistência material, de que são exemplos uma declaração que não foi emitida, uma assembleia geral que não se realizou ou uma deliberação que não foi tomada³¹.

25 O que se segue atualiza o que escrevi em FERREIRA DE ALMEIDA (1992), pp. 426 ss.

26 ADIDA-CANAC (2004), com referências bibliográficas e jurisprudenciais e sustentando o A. a autonomia e o interesse prático da noção.

27 DIEZ-PICAZO (2007), p. 562; BELTRÁN DE HEREDIA (2009), p. 27.

28 Cfr. CARVALHO FERNANDES (2010), pp. 487 ss, com indicação das opiniões a favor e contra. No anteprojeto do Código Civil, ALARCÃO (1959), pp. 200 s, entendeu não tomar partido sobre a *vexata quaestio* da inexistência.

29 Cfr. FILANTI (1983); GALGANO, *Il contratto*, cit., pp. 287 ss.

30 PAIS DE VASCONCELOS (2008), pp. 737 s, distingue três modos de inexistência: ôntica, qualificativa e por imposição legal.

31 CASSIANO SANTOS & DUARTE FONSECA (2012), pp. 49 ss.

2.^a Se o facto não for uma ação. Assim, na coação física e na falta de consciência da emissão de sinais (artigo 246.º).

3.^a Se a ação, sendo exterior ao direito ou ao âmbito negocial, não tiver a natureza de ato ou de negócio jurídico. Assim, nas chamadas declarações não sérias (artigo 245.º)³², na falta de consciência da natureza negocial da declaração³³ (artigo 246.º) e nos donativos conformes aos usos sociais (artigo 940.º, n.º 2, 2.^a parte), pelos quais o direito se desinteressa, deixando o campo aberto para a aplicação de regras sociais exteriores ao direito³⁴.

4.^a Se o ato for ininteligível ou contraditório, com deficiência significativa irreduzível e insolúvel, insuscetível de recuperação por interpretação, com todas as suas dimensões, de tal modo que não se possa apurar um significado em função do qual se estabeleçam os efeitos.

5.^a Se o negócio unilateral ou o contrato sofrerem de insuficiência estrutural extrema ou de contradição sistemática, sem salvação possível por via interpretativa ou integrativa.

Para (pretensos) contratos, são exemplos de insuficiência sintagmática ou de exterioridade paradigmática³⁵, além do dissenso, que é o caso mais relevante e mais referido, a emissão de uma só declaração contratual, a imputação a uma entidade não personalizável, o anonimato de um dos declarantes³⁶, a omissão insuprível do objeto (salvo em contratos familiares) ou de uma circunstância essencial, a indicação de uma função (eficiente) sem valor jurídico (amar, persuadir). Contradição sistemática verifica-se quando haja incompatibilidade inultrapassável entre elementos, como, por exemplo, a coincidência de partes ou o casamento de pessoa física com uma pessoa jurídica ou com um animal. Em todas estas hipóteses, poderá haver um ato jurídico ou até dois, mas não há contrato, o “contrato” é inexistente.

Nas situações 2.^a a 5.^a, a inexistência deriva de o facto não reunir os requisitos mínimos para preencher, respetivamente, os conceitos de ato, de ato jurídico, de declaração jurídica, de negócio jurídico³⁷ ou de contrato, sem que,

32 Isto é, aquelas que se projetam apenas no imaginário artístico, pedagógico, lúdico ou metafórico, alheio ao universo sociojurídico.

33 No sentido de que são inexistentes as cláusulas contratuais gerais excluídas dos contratos singulares, PAIS DE VASCONCELOS (2008), pp. 657 s. Parece-me melhor incluir esta situação no âmbito da ineficácia *stricto sensu* (das cláusulas gerais).

34 Cfr. FERREIRA DE ALMEIDA (2013b), pp. 16 ss.

35 Cfr. FERREIRA DE ALMEIDA (2016), pp. 17 s.

36 O acórdão do STJ de 09.03.2004 (relator Silva Salazar) julgou inexistente um contrato de arrendamento por não haver arrendatário, visto que a pessoa indicada como tal não atribuíra poderes de representação a quem interveio no contrato como seu pretenso representante. A qualificação da situação como de inexistência é discutível, porque a lei a considera como de ineficácia (artigo 268.º, n.º 1).

37 ANDRADE (1964), p. 415, refere-se à inexistência do negócio jurídico por não se verificarem, ao menos, os elementos do conceito.

para tal concluir, seja necessário preceito legal que expressamente proceda a tal qualificação. Não se trata pois de uma valoração como aquela que está subjacente ao juízo de invalidade, mas de pura verificação de não pertença a uma classe de factos ou de atos.

Mas há também hipóteses em que a inexistência está consagrada em disposição de lei. Assim sucede no contrato de casamento que é inexistente nas situações previstas no artigo 1628.^{o38} Os casos da alínea c) – falta de declaração de um ou ambos os nubentes – e porventura da alínea d) – participação de procurador sem poderes – são congruentes com os critérios essenciais antes enunciados, em particular, a insuficiência estrutural extrema na formação do contrato. Já os fundamentos das alíneas a) e b) – incompetência funcional da pessoa perante quem o ato é celebrado e omissão de homologação de casamento urgente – só se compreendem por resquício histórico (francês) da inexistência por ausência de oficial do estado civil. Nas alíneas c) e d), falta o mínimo para que haja acordo, ainda que viciado; nas outras duas alíneas, pressupõe-se a existência de acordo entre os nubentes, embora sem a intervenção de oficial público nos termos da lei. Trata-se portanto de insuficiência de forma, que o sistema jurídico comina geralmente com nulidade³⁹. Nestes casos, como noutros, a qualificação legal de inexistência não passa de ficção para justificar a ineficácia radical.

A inexistência do casamento por ser contraído por duas pessoas do mesmo sexo foi entretanto eliminada, o que demonstra a relatividade dos valores subjacentes aos fundamentos de ineficácia, mesmo quanto à mais extrema das suas modalidades, que é a inexistência.

II. O artigo 1630.^o estabelece o *regime da inexistência* do casamento, dispondo que (n.^o 1) este não produz qualquer efeito jurídico (e nem sequer é havido como putativo) e que (n.^o 2) a inexistência pode ser invocada por qualquer pessoa, a todo o tempo, independentemente de declaração judicial⁴⁰. Este regime tem sido generalizado para outros atos inexistentes, embora por vezes

38 Outros preceitos legais que se referem a inexistência: artigo 354.^o, alínea c) (confissão de facto notoriamente inexistente); artigo 2037.^o, n.^o 1 (devolução da sucessão ao indigno havida como inexistente); Código do Registo Civil, artigos 85.^o e 86.^o, Código do Registo Predial, artigos 14.^o e 15.^o (inexistência dos registos civil e predial); artigo 2.^o do Decreto-Lei n.^o 40/77, de 29 de janeiro (inexistência dos afastamentos de trabalhadores ocorridos entre 25 de abril de 1974 e 25 de abril de 1976, desde que não tenham sido observadas as disposições à data do afastamento sobre a resolução do contrato de trabalho ou tenham ocorrido por motivos políticos ou ideológicos); Constituição, artigos 137.^o, 140.^o, n.^o 2 (inexistência de atos normativos por falta de promulgação ou de referenda).

39 Note-se porém que para o casamento civil a lei não prevê a nulidade, reduzindo as invalidades à inexistência e à anulabilidade.

40 Cfr. com a mesma redação (salvo naturalmente a referência a casamento putativo) o artigo 15.^o do Código do Registo Predial.

se usem expressões mais fortes: que o ato inexistente é nada, que está fora do direito, não se podendo sequer equacionar a produção de efeitos⁴¹; “o facto nunca teve vida, não nasceu sequer – nunca existiu”⁴².

Ora não é bem assim. Tal como o ato nulo, o ato inexistente não produz os efeitos correspondentes ao significado que alguém lhe pretenda atribuir, mas pode produzir efeitos derivados da sua própria inexistência, designadamente a obrigação de restituir o que tiver sido prestado e responsabilidade civil imputável a quem culposamente causar aparência de um ato jurídico (cfr. artigos 245.º, n.º 2, e 246.º, 2.ª parte). Além disso, a invocação por qualquer pessoa e a desnecessidade de declaração judicial devem ser manejadas com cautela, porque, na prática, a inexistência há de ser conhecida oficiosamente⁴³ ou invocada por quem tenha interesse (Código de Processo Civil, artigo 30.º), recaindo sobre o interessado o ónus da prova da inexistência.

III. *Inexistência e nulidade*. Perguntar-se-á então se vale a pena diferenciar inexistência de nulidade, se, como entendem os seus detratores, a distinção não será supérflua⁴⁴, por a inexistência não ter autonomia de regime face à nulidade⁴⁵.

Em minha opinião, a diferença é de manter por razões conceptuais e de regime.

A invalidade afeta atos que se verificaram e que não sofrem de insuficiência ou de ininteligibilidade irreduzíveis. O chamado negócio jurídico inválido é um ato perceptível, porque corresponde, no mínimo, a um modelo estrutural de negócio jurídico e não encerra contradições inultrapassáveis.

A inexistência, material ou jurídica, constitui o grau máximo da ineficácia, é um “vício superior e mais grave do que a nulidade”⁴⁶, uma “supernulidade”⁴⁷, proveniente de omissão do facto referido ou da sua inadmissível qualificação como ato ou como negócio jurídico.

São pois justificadas as soluções no sentido de: 1.º) os factos inexistentes não serem suscetíveis de sanção por conversão ou redução; 2.º) não lhes ser aplicável o regime da inoponibilidade a terceiros prevista no artigo 291.º; e 3.º) o

41 OLIVEIRA ASCENSÃO (2003), p. 373.

42 Acórdão do STJ de 12.12.1995 (relator Lopes Pinto), num caso em que se discutia a inexistência de uma assembleia geral.

43 No sumário do acórdão do STJ de 12.12.1995, já citado, pode ler-se: “O conhecimento [da inexistência] pode impor-se no discorrer de uma argumentação e nessa altura há que a declarar.”

44 HÖRSTER (1992), p. 518.

45 MENEZES CORDEIRO (2014), pp. 927 ss.

46 Sumário do acórdão do STJ de 12.12.1995, já citado.

47 ADIDA-CANAC (2004), n.º 10.

pretensão adquirente não beneficiar do estatuto de possuidor, em qualquer das suas modalidades (artigos 1259.º e seguintes) designadamente para o efeito de usucapião⁴⁸.

Não deve porém ocultar-se que a distinção entre inexistência e invalidez é complexa e subtil. Tanto a invalidez como a inexistência geram ineficácia. A inexistência também é invalidez e, de certo modo, a nulidade também é inexistência, porque o chamado negócio nulo não é, em rigor, negócio jurídico.

4. Caracterização do “contrato nulo”

I. *Em geral.* Um (dito) contrato que não valha como tal, se não é válido enquanto contrato, não pode ser qualificado como contrato. As expressões “negócio jurídico nulo” e “contrato nulo” (ainda que convenientes por simplificação e universalmente usadas) envolvem uma contradição⁴⁹ semântica, mas não há na linguagem jurídica expressão que evite essa contradição.

O contrato, como todos os atos performativos, caracteriza-se pela conformidade dos efeitos com o seu significado. Por isso⁵⁰, se, por inadequação dos sinais emitidos (insuficiência de forma) ou antijuridicidade do comportamento (vício na formação, ilicitude do conteúdo), não se produzirem os efeitos próprios de um contrato ou de um tipo contratual, pode haver um ato socialmente relevante e até juridicamente atendível, mas não há contrato. Se forem outros os efeitos, será outra a natureza do ato.

O chamado contrato nulo é um ato, semelhante a contrato, que é nulo por não observar os requisitos de validade do contrato. Contrato nulo é uma fórmula aberta⁵¹ e multiforme, cujos efeitos jurídicos são variáveis como variável é o fundamento da eficácia que, apesar de tudo, produz, porque não se pode ignorar a sua existência e a correspondente alteração fáctica.

Os negócios jurídicos anuláveis não têm neste aspeto um tratamento jurídico independente. Se forem convalidados (pelo decurso do tempo, por confirmação, por redução ou conversão), tudo se passa como se fossem válidos; se forem anulados, tudo se passa como se fossem nulos (cfr. artigos 289.º e seguintes).

48 Cfr., com algumas variações, MOTA PINTO (2005), p. 619; CARVALHO FERNANDES (2010), p. 490; PAIS DE VASCONCELOS (2008), p. 739.

49 ATYAH (1997), p. 46. A opinião de que os negócios jurídicos nulos não são negócios jurídicos é compartilhada por autores clássicos como ZITELMANN (1879), pp. 86 s., e MOREIRA (1907), pp. 510 s. A opinião contrária raramente resulta da discussão direta do problema, resulta antes da contraposição entre inexistência e nulidade, pressupondo nesta um mínimo de reconhecimento legal.

50 Mas também se outro for o fundamento da eficácia negocial: vontade, confiança, responsabilidade.

51 FILANTI (1983), pp. 83 ss.

Os efeitos dos contratos nulos ou anulados não correspondem ao significado do ato, que teriam se fossem válidos. A consequência primária e fundamental da invalidade é negativa, ainda que não explícita na lei portuguesa: o ato fica sem efeito (*quod nullum est nullum producit effectum*), isto é, com rigor, não produz, nem no futuro nem no passado, os efeitos que produziria se fosse válido.

Mas a declaração ou o reconhecimento da invalidade desencadeiam também consequências positivas, que são as necessárias para corrigir atos sequentes ou até para os manter na medida exigida para compensar expectativas.

O principal conjunto de atos de correção consiste em reconstituir, tanto quanto possível, a situação anterior ao ato, através da restituição do que foi entregue⁵² ou do valor correspondente (artigo 289.º, n.º 1).

Mas nem sempre a invalidade provoca, ou provoca apenas, consequências negativas e de reconstituição. Nalgumas situações, as consequências são atenuadas ou mesmo semelhantes às que o contrato teria se fosse válido. Por exemplo, a posse titulada produz os mesmos efeitos independentemente de o ato que a gerou ser válido ou inválido [artigos 1259.º, 1294.º e 1298.º, alínea a)]. É essa porventura a razão pela qual continua a haver quem sustente, ou subentenda, que o contrato inválido é ainda contrato.

A atenuação dos efeitos restitutivos da invalidade verifica-se, em relação a todos os negócios, por aplicação do regime jurídico da posse (artigo 289.º, n.º 3) e, em relação aos negócios gratuitos, por limitação da restituição à medida do enriquecimento (artigo 289.º, n.º 2). A atenuação é ainda mais forte quando se verifique a exceção que, dentro de certos limites, protege terceiros de boa fé da ineficácia dos negócios inválidos (Código Civil, artigo 291.º⁵³; Código dos Valores Mobiliários, artigo 58.º). Este regime não atinge porém nunca o ponto de conferir, entre as partes, efeitos equivalentes aos que teria o negócio jurídico válido homólogo.

II. Para ser completo o teste da asserção de que os contratos inválidos não são contratos deve ainda passar pelo crivo dos regimes de *alguns tipos contratuais* duradouros (de sociedade comercial, de trabalho e de seguro), que, sob este aspeto, parecem mais problemáticos⁵⁴.

52 A frase “o que tiver sido prestado” deve ser interpretada de modo a abranger qualquer atribuição, mesmo que não corresponda a uma prestação – cfr. MÚRIAS (2008), pp. 797 ss –, o que sucede, designadamente, com as entregas iniciais em contratos reais quanto à constituição.

53 Ver, por último, SOTTOMAYOR (2010).

54 Seguindo a orientação, que venho adotando, de construir a teoria geral dos contratos também com base no regime dos tipos contratuais.

Em relação ao *contrato de sociedade comercial*, o regime especial das invalidades incide sobre vários pontos, que, para proteção da subsistência da pessoa jurídica, do património autónomo ou de terceiros, 1.º restringem os fundamentos da nulidade (Código das Sociedades Comerciais, artigo 42.º), 2.º atenuam os efeitos da invalidade, substituindo a mera restituição pela liquidação patrimonial (artigo 52.º, n.º 1) e pela exoneração (artigo 45.º), e 3.º permitem a sanção de nulidades (artigos 42.º, n.º 2, e 43.º, n.º 3).

O aspeto mais controverso é o enquadramento jurídico da situação prevista no artigo 36.º, n.º 2, do mesmo Código⁵⁵. Para o que agora importa as teses mais significativas são três: 1.ª nulidade do contrato de sociedade comercial por falta de forma⁵⁶ com aplicação do regime da sociedade civil⁵⁷; 2.ª sociedade civil por conversão *ope legis*⁵⁸; 3.ª contrato associativo *sui generis* válido⁵⁹.

De nenhuma daquelas especialidades de regime – restrição dos fundamentos, atenuação dos efeitos, sanção de invalidades – e de nenhuma destas construções resulta que o ato em causa, quando seja inválido, conduza, com o registo, à constituição de uma sociedade comercial, como é próprio do contrato de sociedade comercial. Não fica pois em crise o princípio de que o (chamado) contrato de sociedade comercial inválido não produz os efeitos típicos deste contrato, pelo que não se pode, em rigor, qualificar como contrato de sociedade comercial, embora desencadeie alguns (limitados) efeitos secundários semelhantes ao que este produz.

“O *contrato de trabalho* declarado nulo ou anulado produz efeitos como válido em relação ao tempo em que seja executado” (Código do Trabalho, artigo 122.º, n.º 1). Para esta regra⁶⁰, à primeira vista anómala, porque a invalidade não perturba os efeitos nucleares do contrato – o trabalho e o salário – têm-se procurado diferentes enquadramentos, como sejam a ficção de validade e a recondução a uma relação contratual de facto.

Mas a regra não é assim tão estranha. Não havendo modo de devolver o trabalho já feito, a lei consagra um meio expedito de concretização da regra de restituição por equivalente (artigo 289.º, n.º 1, *in fine*), presumindo-se, *juris et de jure*, que o salário pago, que, segundo o regime geral, haveria de ser restituído,

55 “Se for acordada a constituição de uma sociedade comercial, mas, antes da celebração do contrato de sociedade, os sócios iniciarem a sua atividade, são aplicáveis às relações estabelecidas entre eles com terceiros as disposições sobre sociedades civis.”

56 PITA (2004), conclusão na p. 313.

57 COUTINHO DE ABREU (2011), pp. 119 s.

58 PINTO FURTADO (2004), p. 211; acórdão do STJ de 08.11.2005 (relator Azevedo Ramos).

59 CASTELO (2011), pp. 78, 92.

60 Que se deve entender como aplicável também a casos de inexistência e de ineficácia: MENEZES LEITÃO (2014), p. 252.

corresponde ao valor da prestação de trabalho⁶¹. Quanto ao restante, verificam-se, sem alteração, os efeitos comuns da invalidade: cessam as obrigações de prestação de trabalho e da sua retribuição, assim como todos os outros direitos e deveres recíprocos próprios das relações contratuais de trabalho.

Este regime legal não convalida sequer o contrato durante o tempo da sua execução⁶². O chamado contrato de trabalho nulo ou anulado não é (*não foi*) afinal um contrato de trabalho.

Quanto ao *contrato de seguro*. O contrato, no qual alguém participe como segurador sem estar autorizado a exercer a atividade seguradora em Portugal no âmbito do respetivo ramo, é nulo, mas, se a contraparte ignorar esta situação⁶³, a pessoa que se apresenta como segurador autorizado não se exime do cumprimento das obrigações que para ele decorreriam do contrato ou da lei, caso o negócio fosse válido⁶⁴.

A doutrina portuguesa maioritária pronuncia-se no sentido de a participação de um segurador autorizado ser elemento necessário para a tipificação do contrato de seguro⁶⁵. O contrato de seguro é um contrato próprio em função da estrita conexão de uma estrutura empresarial especializada e sujeita a supervisão pública com a aptidão para proceder ao cálculo atuarial e para assegurar a capacidade financeira que esta categoria de negócios exige. Daí a nulidade do contrato em que não se verifique a participação de uma empresa com estas características.

Mas há quem entenda, pelo contrário, que a empresa não é elemento necessário para qualificar um contrato como de seguro, pelo que um contrato não deixa de ser de seguro se o segurador não for uma empresa seguradora⁶⁶.

61 Próximo, MONTEIRO FERNANDES (2014), p. 314, observando que o mecanismo civil também conduziria à neutralização dos efeitos da nulidade no período da execução do trabalho. ROMANO MARTINEZ (2015), p. 420, qualifica a situação como “fissão de validade” e considera, sem mais, que não se aplica o regime do Código Civil.

62 RAMALHO (2014), p. 188.

63 Para que haja boa fé do tomador, a hipótese mais provável é que a outra parte esteja autorizada a celebrar contratos de seguro noutra Estado ou em ramo diferente daquele em que o contrato se insere.

64 Lei do Contrato de Seguro, artigo 16.º, cujo teor é o seguinte:

“1 – O segurador deve estar legalmente autorizado a exercer a atividade seguradora em Portugal, no âmbito do ramo em que atua, nos termos do regime jurídico de acesso e exercício da atividade seguradora. 2 – Sem prejuízo de outras sanções aplicáveis, a violação do disposto no número anterior gera nulidade do contrato, mas não exime aquele que aceitou cobrir o risco de outrem do cumprimento das obrigações que para ele decorreriam do contrato ou da lei caso o negócio fosse válido, salvo havendo má fé da contraparte.”

65 Ver, entre outros, ENGRÁCIA ANTUNES (2009a), pp. 683 s; PAIS DE VASCONCELOS (2009), p. 159 (como índice do tipo); PINTO DUARTE (2000), p. 68, nota 183 (referindo-se a uma opinião comum e citando bibliografia). Em obras estrangeiras, CHULIÁ (2007), pp. 1087, 1093; DE LORENZI (2008), pp. 39, 42, 44, 65 ss; HOFMAN (1991), pp. 31 ss.

66 LIMA REGO (2010), pp. 303 ss.

O universo dos contratos de seguro seria mais amplo do que o dos contratos de seguro admissíveis na nossa ordem jurídica⁶⁷. Em consequência, os contratos de seguro celebrados por entidades não autorizadas, embora sejam nulos, são contratos de seguro⁶⁸.

Vejamos, antes de mais, o que resulta do citado artigo 16.º, combinado com as regras gerais sobre nulidade. Esta é de conhecimento oficioso pelo tribunal e pode ser invocada a todo o tempo pelo tomador e por qualquer outro interessado, salvo pela pessoa que surge como segurador⁶⁹.

Se, ao tempo da invocação, não se tiver verificado sinistro, cessam total e retroativamente todos os efeitos do contrato, com devolução do prémio e desresponsabilização da contraparte por sinistro futuro.

Se, ao tempo da invocação eficaz de nulidade, se tiver verificado um sinistro, o tomador do seguro de boa fé (isto é, que ignorava a falta de autorização para a contraparte celebrar contratos de seguro ou contratos de seguro no âmbito daquele ramo) tem direito a receber da contraparte prestação igual à que resultaria de um contrato de seguro válido⁷⁰.

Sendo o “contrato” nulo por incapacidade de gozo⁷¹ de quem se apresenta como segurador, como se justificam estes efeitos? Em minha opinião, a justificação não reside em que seja ainda um contrato de seguro.

Note-se, em primeiro lugar, que o destino normal deste contrato nulo é cessar sem produzir nenhum dos efeitos característicos de um contrato de seguro. Só gera um dos efeitos semelhantes a um contrato de seguro válido (a obrigação do pretense segurador) se se verificarem três requisitos cumulativos: que

67 LIMA REGO (2010), p. 39.

68 LIMA REGO (2010), p. 40; P. ROMANO MARTINEZ, anotação em *Lei do contrato de seguro anotada* (2008), p. 78.

69 Não se trata pois de uma invalidez mista – como diz MENEZES CORDEIRO (2013), p. 472 – no sentido de que só pode ser invocada por uma das partes. Como o exercício autorizado da atividade seguradora é de ordem pública – EDUARDA RIBEIRO, em *Lei do contrato de seguro anotada* (2008), p. 74 – não se deve excluir o conhecimento oficioso e devem incluir-se no âmbito dos interessados o Ministério Público, a Autoridade de Supervisão dos Seguros e qualquer segurador autorizado para o mesmo ramo.

70 É duvidoso se, neste caso, o pretense segurador tem ou não direito a reter os prémios pagos ou a receber o pagamento dos prémios vencidos e não pagos. Sendo a lei omissa, inclino-me para a solução negativa, aplicando-se os efeitos normais da nulidade, só excecionados nos restritos termos fixados pela lei. O artigo 16.º, n.º 2, funcionará assim como uma sanção civil, qualificação usada por EDUARDA RIBEIRO, *Lei do contrato de seguro anotada* (2008), p. 77.

71 Contra, MENEZES CORDEIRO (2013), pp. 472 s, com o argumento de que as incapacidades protegem os próprios, não terceiros. Mas tal só é exato para as incapacidades de exercício. Há incapacidades de gozo – entendidas como fatores excludentes do âmbito das situações jurídicas de que uma pessoa, física ou jurídica, pode ser titular, ativo ou passivo – que são estabelecidas por lei no interesse de terceiros ou também no interesse de terceiros; exemplos: incapacidade para testar, para casar e para celebrar convenção antenupcial, nos termos do artigo 1720.º, n.º 1, alínea a) (tendo em conta interesses de potenciais herdeiros), incapacidade de associações e de sociedades (tendo em conta interesses dos associados, dos sócios e dos credores).

a invalidade não tenha sido invocada, que a parte em posição correspondente ao tomador do seguro esteja de boa fé e que se tenha verificado um sinistro. De outro modo, a incapacidade de gozo tem consequências iguais à de um comum “contrato nulo”. Se se entender que é um contrato de seguro, será um estranho contrato de seguro, cuja qualificação como tal depende da concretização do risco, próximo pois de um contrato de jogo ou de um contrato de risco atípico e proibido. Por isso, o n.º (V) do preâmbulo da lei a que pertence o citado artigo 16.º se refere a “pretensos contratos de seguro”⁷².

A ideia de distinguir os planos de qualificação e da admissibilidade do contrato de seguro não é metodologicamente correta, porque, fora do direito comparado, da história e da filosofia do direito, as qualificações fazem-se com base no regime vigente numa determinada ordem jurídica. O exemplo do “seguro de caso único”⁷³ é neutro, porque, se o contrato for celebrado por empresa seguradora autorizada, o contrato é válido seja como contrato de seguro seja como contrato de aposta⁷⁴.

A mais ardorosa defensora da qualificação como contrato de seguro do contrato nulo por falta de autorização para o exercício da atividade respetiva vem aliás a matizar (ou a contrariar?) a sua posição, admitindo afinal que o ato se possa qualificar como contrato misto com elementos do contrato de seguro⁷⁵.

Parece-me que a nulidade do “contrato de seguro” descaracteriza o tipo contratual e até a qualificação como contrato. Os limitados efeitos que o ato produz justificam-se pela assimilação ao abuso de direito em que incorreria o pretense segurador que viesse a invocar a nulidade do contrato. Um tal comportamento seria *venire contra factum proprium*, contrário aos ditames da boa fé. Daí que não possa exercer o que seria o direito de, em abstrato, invocar a sua própria incapacidade de gozo. Daí também a exigência da boa fé (subjéctiva) do tomador do seguro, porquanto, em geral, a má fé de quem confia neutraliza o instituto do abuso de direito.

A circunstância de o regime constar da lei não impede esta justificação e esta qualificação. O artigo 334.º Código Civil é uma cláusula geral proibitiva do abuso de direito, por excesso manifesto dos limites impostos pela boa fé, pelos bons costumes e/ou pelo fim social ou económico do direito. Mas não tem, não

72 Alguns destes argumentos coincidem com os que são aduzidos por ENGRÁCIA ANTUNES (2009b), pp. 273 ss.

73 LIMA REGO (2010), p. 304.

74 Para GALVÃO TELLES, I. (1950), pp. 80 s, o seguro isolado seria aposta. A qualificação como aposta (válida, porque admitida pela Lei do Contrato de Seguro, que não o proíbe) não fica descaracterizada se o contrato for real quanto à formação, pela entrega de um prémio inicial e único; cfr. FERREIRA DE ALMEIDA (2013b), pp. 237 e 270.

75 LIMA REGO (2010), pp. 307 s.

pretende ter, exclusividade normativa reguladora da proibição do excesso no exercício do direito.

A diferença entre o abuso de direito fundado nesta cláusula geral e o abuso de direito fundado numa lei específica consiste em que, nesta segunda hipótese, o comportamento abusivo está descrito na própria lei. Por isso, não há necessidade de recorrer a parâmetros extrajurídicos. Os requisitos e efeitos são os que constarem da lei, sem que outros, positivos ou negativos, sejam relevantes.

No caso, por força da lei, a confiança justificada basta-se com a ignorância da incapacidade de gozo de quem se apresenta como segurador e o investimento de confiança basta-se com a expectativa de que o risco está coberto pelo contrato⁷⁶. O efeito é a obrigação de efetuar a prestação que caberia ao segurador autorizado em caso de sinistro, que equivale exatamente ao efeito fixado pelo citado artigo 334.º: não poder exercer o direito (melhor, a exceção) de invocar o que seria o efeito da nulidade quanto à obrigação de quem agiu abusivamente.

III. Confirma-se pois que a expressão “*contratos nulos*” (ou anulados) é um modo simplificado de referir atos jurídicos que, tendo alguma semelhança estrutural com contratos, *não valem como contratos* e, por isso, não produzem efeitos iguais a contratos ou aos tipos contratuais a que se assemelham. Os efeitos que produzem são efémeros e têm um fundamento jurídico diferente dos contratos válidos homólogos.

Os efeitos comuns da nulidade – a cessação e a restituição – são inversos dos efeitos próprios dos contratos. Nos tipos contratuais analisados, esses efeitos ou são equivalentes da restituição (no contrato de trabalho) ou derivam de requalificação ou de conversão (no contrato de sociedade comercial nulo por razão formal) ou justificam-se pela correção do abuso de direito (no contrato de seguro nulo por falta de autorização para o exercício da atividade seguradora).

Os contratos inválidos nunca têm os efeitos próprios dos contratos em geral ou do tipo contratual correspondente. Não são portanto contratos ou não são contratos do tipo que teria o contrato válido correspondente.

Mas não são atos inexistentes, porque, para serem nulos ou anulados, é necessário que tenham ocorrido e que sejam atos jurídicos com conteúdo compreensível e isentos de insuficiência estrutural ou de contradição sistemática extremas.

76 Situação de confiança justificada, investimento de confiança e boa fé de quem confiou são os requisitos que pacificamente a doutrina e a jurisprudência exigem para aplicar o instituto do abuso de direito por violação da boa fé; por todos, MENEZES CORDEIRO (2005), pp. 409 ss.

5. Ineficácia em sentido restrito

I. *Conceito e categorias*. Em sentido amplo – não produção de efeitos – a ineficácia inclui a invalidade (ato nulo ou anulado), a inexistência e a ineficácia em sentido estrito.

A dicotomia invalidade/ineficácia *stricto sensu* é aceite de modo quase unânime pela doutrina portuguesa⁷⁷. Em textos legais, a alternativa é expressa, por exemplo, nos artigos 19.º, n.º 1, 36.º, n.º 1, 65.º, n.º 2, e 1717.º do Código Civil. Outros preceitos legais são sugestivos da referência à ineficácia em sentido restrito (por exemplo, ineficácia da declaração negocial, artigo 224.º, n.º 3, e 226.º, n.º 2, ineficácia da proposta, artigo 231.º, n.º 2, ineficácia da convenção antenupcial por caducidade, artigo 1705.º, 1).

O critério distintivo subjacente consiste em que, na invalidade, a ineficácia deriva do desvalor jurídico (ou vício) reportado a um elemento ou requisito intrínseco à estrutura ou à formação do ato, enquanto a ineficácia em sentido estrito resulta de um facto extrínseco ao ato e, por isso, não valorativo deste⁷⁸.

A ineficácia em sentido estrito verifica-se se um ato existente e válido não produzir imediatamente (ineficácia originária) ou deixar de produzir (ineficácia subsequente) a totalidade ou parte dos seus efeitos⁷⁹.

A ineficácia *stricto sensu* não é uma sanção nem um efeito sancionatório; é apenas a consequência de conformidade com a autonomia privada ou de desconformidade não valorativa com certas regras legais.

Não há regras gerais que se apliquem à mera ineficácia⁸⁰, porque as situações de ineficácia *stricto sensu* são diversas e heterogêneas, como demonstram as classificações e os exemplos que se seguem.

A mera ineficácia é autónoma, quando decorre da autonomia privada. É heterónoma, quando a eficácia de ato válido depende (ou a ineficácia resulta) apenas da lei ou de ato jurídico previsto na lei.

77 A exceção é ALARCÃO (1983), pp. 609 ss, para quem a invalidade só é útil como expediente expositivo e sistemático, mas não se impõe enquanto categoria dogmática e científica fundada num critério distintivo da ineficácia estrita (pp. 624 ss). Para DIEZ-PICAZO (2007), pp. 563 s, a distinção entre invalidade e ineficácia não parece útil. O A. do presente artigo espera que o que se segue demonstre que a distinção é útil e baseada em critérios rigorosos.

78 Era este o critério de BETTI (1950), pp. 456 s, e continua a ser o critério dominante na doutrina portuguesa; ver, entre outros, GALVÃO TELLES, I. (1962), pp. 378 s; MOTA PINTO (2005), p. 615; CARVALHO FERNANDES (2010), p. 542.

79 Esta oposição corresponde aproximadamente à distinção alemã entre ineficácia suspensa (*schwebende Unwirksamkeit*) e eficácia suspensa (*schwebende Wirksamkeit*); cfr. WOLF & NEUNER (2012), pp. 666 ss.

80 No anteprojeto do Código Civil, decidiu-se não regular, em geral, a simples ineficácia ou ineficácia *stricto sensu*, remetendo-se a sua regulação parcelar para cada uma das várias formas que pode revestir: ALARCÃO (1959), p. 202.

II. A *ineficácia é autónoma* (ou endógena) quando tem como fonte uma cláusula negocial que se revela no conteúdo do próprio ato⁸¹.

A ineficácia autónoma pode aplicar-se *ipso facto* ou depender da prática de um ato de uma ou de ambas as partes ou de terceiro.

A ineficácia autónoma aplicável *ipso facto* pode ser originária ou subsequente. É originária, se deriva de termo inicial, de condição ou de outra circunstância de eventualidade com efeito suspensivo. É subsequente, se deriva de condição resolutiva, de termo final ou de outra cláusula que determine a caducidade total ou parcial dos efeitos do ato.

A ineficácia autónoma pode também ser potestativa, quando se funda no contrato mas depende de um ato ulterior livre (por exemplo, resolução ou denúncia do contrato nos termos de cláusula contratual). Estes atos determinam sempre ineficácia subsequente.

A ineficácia autónoma é tendencialmente absoluta, no sentido de que afeta quaisquer pessoas – em primeiro lugar as partes, mas também os seus sucessores e terceiros.

III. A *ineficácia é heterónoma* (ou exógena), se tiver como fonte a lei. É geralmente subsequente e pode funcionar *ipso iure* (por exemplo, moratória legal) ou ser potestativa (por exemplo, resolução por incumprimento; revogação pelo doador do contrato de doação, nos termos dos artigos 970.º a 979.º; revogação unilateral da declaração contratual do consumidor no âmbito do direito de arrendimento⁸²).

A ineficácia heterónoma pode ser absoluta, reportar-se apenas aos efeitos entre as partes ou restringir os efeitos em relação a terceiros.

A ineficácia é absoluta, designadamente, em situações de falta de homologação ou autorização, de caducidade *ex lege*⁸³, como sanção por violação de um dever legal⁸⁴ e por falta dos requisitos para inserção de cláusulas contratuais gerais em contratos singulares⁸⁵.

81 Não há contradição entre o carácter extrínseco do facto determinante da eficácia *stricto sensu* e a existência de uma categoria de mera ineficácia autónoma. O facto gerador da ineficácia estrita é sempre extrínseco em relação ao ato, mas aquele facto pode estar previsto no próprio ato, que funciona então como fonte autónoma de ineficácia. É o que sucede na condição e no termo que, sendo cláusulas válidas, determinam a ineficácia do ato se se verificarem os factos extrínsecos previstos nessas cláusulas. A aparente dificuldade ou incongruência resultam de tanto a condição como o termo terem o duplo sentido de cláusulas e de factos previstos nessas cláusulas.

82 Neste sentido, FERREIRA DE ALMEIDA (2005b), p. 113, mas a natureza do ato do consumidor é controversa.

83 Exemplo: caducidade da proposta nos casos do artigos 228.º, n.º 1, alíneas c) e b), e 231.º, 2.

84 Por exemplo, inibição dos direitos de voto e a dividendos por violação do dever de lançamento de uma oferta pública de aquisição (Código dos Valores Mobiliários, artigo 192.º).

85 Decreto-Lei n.º 446/85, de 25 de outubro, artigos 5.º, 6.º e 8.º

A ineficácia tem apenas efeitos entre as partes nas situações de exercício abusivo de um direito (artigo 334.º).

A ineficácia limitada a terceiros diz-se inoponibilidade, quando é originária, em consequência, por exemplo, da omissão de registo declarativo⁸⁶ ou de notificação da cessão de crédito, do contrato ou do penhor (artigos 583.º, n.º 1, 424.º, n.º 2, e 681.º) e da prática de atos de disposição pelo insolvente (Código da Insolvência, artigo 81.º, n.º 6) ou pelo titular de bens arrestados (artigo 622.º, n.º 1); diz-se impugnabilidade⁸⁷, quando é superveniente (impugnação pauliana, artigos 610.º e seguintes; redução de liberalidades, artigos 2169.º e seguintes).

6. Referentes da invalidade, da inexistência e da ineficácia *stricto sensu*

A invalidade, porque assente em juízos de valor, não se refere a quaisquer factos, mas apenas a atos jurídicos – negócios jurídicos (contratos e negócios jurídicos unilaterais, como a procuração e o testamento) e a outros atos jurídicos unilaterais declarativos. No direito público, a invalidade reporta-se também sempre a atos jurídicos: a lei, o ato e o regulamento administrativo, a sentença e outros atos processuais do tribunal ou das partes, a obtenção de prova, os registos públicos.

Alguns atos jurídicos (contratos escritos, sentença, registos) criam (ou constam de) um documento. A invalidade pode resultar de ilicitude no seu modo de formação ou de contrariedade do seu conteúdo com o sistema jurídico. Mas a invalidade reporta-se sempre ao ato, que, sendo nulo ou anulado, não produz os seus efeitos típicos. Para o desvalor do texto e do eventual documento os efeitos negativos são reflexos.

E em relação a direitos e outras situações jurídicas, será correto falar em invalidade?

A dúvida pode suscitar-se em especial para os direitos da propriedade industrial, inquirindo designadamente se é correta a expressão “invalidade da marca”. A resposta é, a meu ver, negativa. Os artigos 33.º, n.º 1, e 34.º, n.º 1, do respetivo Código ditam os fundamentos de nulidade e de anulabilidade total ou parcial “[d]as patentes, [d]os modelos de utilidade e [d]os registos”; os arti-

86 Segundo a opinião comum; ver, porém, FERREIRA DE ALMEIDA (2014), pp. 106 s.

87 No sentido próprio e restrito que lhe atribui uma parte da doutrina: OLIVEIRA ASCENSÃO (2003), pp. 403 s; CARVALHO FERNANDES (2010), pp. 544 s, ambos citando obras portuguesas anteriores. Mas impugnar e impugnação são palavras que a lei também usa em sentido muito amplo para referir a contestação da existência, da validade ou da eficácia de um ato (artigos 642.º, n.º 2, 698.º, n.º 2, 765.º, n.º 2, 993.º, n.º 2, 1433.º, n.º 1, 1859.º ss) ou de uma relação jurídica (impugnação da paternidade, artigos 1838.º ss) ou a veracidade de um documento (artigos 368.º, 374.º, n.º 2).

gos 35.º e 36.º regulam os processos e os efeitos da declaração de nulidade e da anulação. Ora, apesar da aparente alternativa legal, a invalidade refere-se sempre a registos, sejam eles de marca, de logótipo, de desenho ou modelo (cfr. artigos 208.º e 209.º) ou de patente, que tem a natureza (de registo) de uma concessão administrativa⁸⁸ atribuída a uma invenção (cfr. artigo 51.º). A referência da invalidade a atos jurídicos não é portanto perturbada neste campo, por ser esta a natureza tanto dos registos como da concessão administrativa.

Em geral, pode dizer-se que a invalidade nunca se reporta diretamente a direitos ou a outras situações jurídicas, mas ao ato de que pretensamente emerge (ver, como exemplo, o artigo 1416.º, que estabelece os fundamentos da “nulidade do título constitutivo da propriedade horizontal” e a “legitimidade para arguir a nulidade do título”). Em rigor, a invalidade tem portanto sempre um ato jurídico como referente.

Bem mais amplos são os referentes possíveis da inexistência. Desde logo, porque a inexistência se pode relacionar com quaisquer factos, sejam ou não atos jurídicos. É assim quanto à ocorrência da condição casual ou de outro evento relevante (por exemplo, o sinistro, no contrato de seguro), quanto à receção efetiva de uma declaração e, em geral, quanto a qualquer facto, ou conhecimento do facto, que seja objeto de prova. Trata-se então sempre de saber se o facto – incluindo um facto natural e um facto humano, seja ou não uma ação – se verificou ou não, isto é, se existiu ou não.

Em relação a atos jurídicos ou pretensamente jurídicos, a inexistência nem sempre se afere pela simples verificação histórica de que não ocorreram. Envolve uma qualificação para apurar, como se viu (n.º 3-I), se o facto que ocorreu reúne os requisitos mínimos para preencher os conceitos de ato, de ato jurídico, de declaração jurídica, de negócio jurídico ou de contrato ou se preenche uma específica previsão legal para ser apodado de inexistente.

A inexistência surge também em relação a pessoas, coisas e documentos, com relevância para a imputação de atos em que se invoca pessoa inexistente, a impossibilidade do objeto e da prestação (artigos 280.º, n.º 1, e 790.º) a (in)idoneidade de coisa futura como objeto negocial e a falsidade de documento (artigos 372.º, 375.º, n.º 2, 376.º, n.º 1).

Finalmente, a inexistência poderia reportar-se a direitos e a outras situações jurídicas, dizendo-se, por exemplo, que um dado direito ou dever de uma concreta pessoa não existem. Mas uma tal inexistência significa afinal um juízo de não relevância jurídica, sem autonomia científica, didática ou prática, porque se dissolve no âmbito da aplicação do direito, em geral, e da decisão de controvérsias jurídicas, em especial.

88 OLIVEIRA ASCENSÃO (1988), pp. 243 ss, 327 ss.

A ineficácia *stricto sensu* tem sido construída e exemplificada com referência a atos jurídicos. Mas, por definição, a ineficácia consiste em excluir ou suspender efeitos cujo conteúdo se preenche com situações jurídicas. Por outro lado, a ineficácia destas pode resultar da aplicação de institutos autónomos em relação aos atos que as criam, como são o abuso e a colisão de direitos (artigos 334.º e 335.º), que comprimem o seu exercício (isto é, a eficácia) em dadas circunstâncias ou na sua plenitude, sem todavia extinguirem o direito ou outra situação jurídica ativa.

Embora se possam também referir a documentos, designadamente a documentos caducados (que, na linguagem comum, impropriamente se dizem “fora do prazo de validade”), os referentes próprios da ineficácia em sentido restrito são portanto os atos jurídicos e as situações jurídicas.

A determinação dos referentes possíveis da invalidade, da inexistência e da ineficácia é relevante para o rigor terminológico (e o direito é feito quase só de palavras). Mas, por isso mesmo, é também importante para delimitar a previsão ou a estatuição de algumas normas, esclarecendo os respetivos regimes jurídicos.

Bibliografia

- ADIDA-CANAC, H. (2004), “Actualité de l’inexistence des actes juridiques”, http://www.courdecassation.fr/publications_26/rapport_annuel_36/rapport_2004_173/deuxieme_partie_tudes_documents_176/tudes_diverses_179/actes_juridiques_6399.htm.
- ALARCÃO, RUI DE (1959), “Invalidade dos negócios jurídicos”, *Boletim do Ministério da Justiça*, n.º 89, pp. 199 ss.
- , (1983), “Sobre a invalidade do negócio jurídico”, *Estudos em Homenagem ao Prof. Doutor J. J. Teixeira Ribeiro*, III, *Jurídica*, Coimbra, pp. 609 ss.
- ALEXANDRE, ISABEL (1998), *Provas ilícitas em processo civil*, Coimbra.
- ANDRADE, MANUEL DE (1964), *Teoria geral da relação jurídica*, II, *Facto jurídico, em especial negócio jurídico*, Coimbra, II.
- ATIYAH, P. S. (1997), *An Introduction to the Law of Contract*, 5.ª ed., Oxford.
- AUSTIN, J. L. (1962), *How to do Things with Words*, Oxford.
- BELTRÁN DE HEREDIA, C. L. (2009), *La nulidad de los contratos*, Valencia.
- BETTI, E. (1950), *Teoria generale del negozio giuridico*, Torino, reimp. Napoli, 1994.
- BRITO, M. HELENA (1999), *A representação nos contratos internacionais*, Coimbra.
- CARVALHO FERNANDES, L. (2010), *Teoria Geral do Direito Civil*, II, 5.ª ed., Lisboa.
- CASSIANO SANTOS & DUARTE FONSECA (2012), “Inexistência e nulidade de deliberações”, *Direito das Sociedades em Revista*, vol. 7, pp. 49 ss.

- CASTELO, HIGINA (2011), “O acordo a que se reporta o artigo 36.º, 2, do Código das Sociedades Comerciais – natureza e validade”, *Estudos em homenagem a CFA*, Coimbra, IV, pp. 51 ss.
- CHULIÁ, F. V. (2007), *Introducción al Derecho Mercantil*, 20.ª ed., Valencia.
- COSTA ANDRADE, M. (1991), *Sobre as proibições de prova em Processo Penal*, Coimbra, reimp. 2006.
- COUTINHO DE ABREU, J. M. (2011), *Curso de direito comercial*, II. *Das Sociedades*, 4.ª ed., Coimbra.
- DE LORENZI, V. (2008), *Contratto di assicurazione. Disciplina giuridica e analisi economica*, Padova.
- DIAS MARQUES, J. (1992), *Noções Elementares de Direito Civil*, 7.ª ed., Lisboa
- DÍEZ-PICAZO, L. (2007), *Fundamentos de derecho civil patrimonial*, I, *Introducción. Teoría del contrato*, 6.ª ed., Cizur Menor.
- DÖRR & SCHMALENBACH (2012), *Vienna Convention on the Law of Treaties. A Commentary*, Berlin, Heidelberg.
- ENGRÁCIA ANTUNES, J. (2009a), *Direito dos Contratos Comerciais*, Coimbra.
- , (2009b), “‘Contrato de seguro e terceiros’. Arguição de provas públicas de doutoramento de Margarida Lima Rego”, *Themis*, n.º 17, pp. 267 ss.
- FERREIRA DE ALMEIDA, C. (1992), *Texto e enunciado na teoria do negócio jurídico*, Coimbra.
- , (2005a), “Registo de valores mobiliários”, *Estudos em memória do Professor Doutor António Marques dos Santos*, Coimbra, vol. I, pp. 873 ss.
- , (2005b), *Direito do Consumo*, Coimbra.
- , (2013a), *Contratos I. Conceito. Fontes. Formação*, 5.ª ed., Coimbra.
- , (2013b), *Contratos III. Contratos de liberalidade, de cooperação e de risco*, 2.ª ed., Coimbra.
- , (2014), *Contratos IV. Funções. Circunstâncias. Interpretação*, Coimbra.
- , (2016), *Contratos II. Conteúdo. Contratos de troca*, 4.ª ed., Coimbra.
- FILANTI, G. (1983), *Inesistenza e nullità*, Napoli.
- FLOUR, AUBERT & SAVAUX (2000), *Les obligations*, 1. *L'acte juridique*, 9.ª ed., Paris.
- GALGANO, F. (2007), *Il contratto*, Padova.
- GALVÃO TELLES, M. (1967), *Eficácia dos tratados na ordem interna portuguesa (condições, termos e limites)*, Lisboa.
- GALVÃO TELLES, I. (1950), “Aspectos comuns aos vários contratos”, *Revista da Faculdade de Direito da Universidade de Lisboa*, separata.
- , (1962), *Manual dos contratos em geral*, 2.ª ed., Lisboa.
- GOMES CANOTILHO, J. J. (2003), *Direito constitucional e teoria da Constituição*, 7.ª ed., Coimbra.
- HOFMAN, E. (1991), *Privatversicherungsrecht*, 3.ª ed., München.

- HÖRSTER, H. E. (1992), *A parte geral do Código Civil português. Teoria Geral do Direito Civil*, Coimbra.
- LASARTE, C. (2006), *Contratos*, 9.ª ed., Madrid, Barcelona.
- LEBRE DE FREITAS, J. (2008), *Código de Processo Civil anotado*, vol. 1.º, 2.ª ed., Coimbra.
- , (2001), *Código de Processo Civil anotado*, vol. 2.º, Coimbra.
- Lei do contrato de seguro anotada* (2008), org. ROMANO MARTINEZ e o., Coimbra.
- LIMA REGO, M. (2010), *Contrato de seguro e terceiros*, Coimbra.
- MEDEIROS, RUI (1999), *A decisão de inconstitucionalidade*, Lisboa.
- MENEZES CORDEIRO, A. (2005), *Tratado de Direito Civil Português*, I, *Parte Geral*, tomo I, 3.ª ed., Coimbra.
- , (2008), “Da ineficácia civil. Reflexões críticas”, *Estudos em Honra do Professor Doutor José de Oliveira Ascensão*, Coimbra, I, pp. 233 ss.
- , (2013), *Direito dos Seguros*, Coimbra.
- , (2014), *Tratado de Direito Civil*, II, *Parte Geral. Negócio jurídico*, 4.ª ed., Coimbra.
- MENEZES LEITÃO, L. (2014), *Direito do Trabalho*, 4.ª ed., Coimbra.
- MONTEIRO FERNANDES, A. (2014), *Direito do Trabalho*, 17.ª ed., Coimbra.
- MOREIRA, GUILHERME (1907), *Instituições de Direito Civil Português*, I, Coimbra.
- MOTA PINTO, C. (2005), *Teoria geral do direito civil*, 4.ª ed. por A. Pinto Monteiro e P. Mota Pinto, Coimbra.
- MÚRIAS, PEDRO (2008), “Um conceito de atribuição para o direito do não cumprimento”, *O Direito*, IV, pp. 797 ss.
- OLIVEIRA ASCENSÃO, J. (1988), *Direito Industrial*, Lisboa.
- , (2003), *Direito civil. Teoria Geral*, II, *Acções e Factos Jurídicos*, 2.ª ed., Coimbra.
- PACHECO DE AMORIM, J. (2008), “A invalidade e a (in)eficácia do contrato administrativo no Código dos Contratos Públicos”, *Estudos de Contratação Pública*, org. Pedro Gonçalves, Coimbra, I, pp. 627 ss.
- PAIS DE VASCONCELOS, P. (2008), *Teoria geral do direito civil*, 5.ª ed., Coimbra.
- , (2009), *Contratos atípicos*, 2.ª ed., Coimbra.
- PINTO DUARTE, R. (2000), *Tipicidade e atipicidade dos contratos*, Coimbra.
- PINTO FURTADO, J. (2004), *Curso de Direito das Sociedades*, 5.ª ed., Coimbra.
- PITA, MANUEL A. (2004), *O regime da sociedade irregular e a integridade do capital social*, Coimbra.
- Principles. Definitions and Model Rules of European Contract Law. Draft Common Frame of Reference (DCFR)*, 2009, Full Edition (org. Von Bar & Clive), Munich, vol. 1
- RAMALHO, M. ROSÁRIO (2014), *Tratado de Direito do Trabalho. Parte II – Situações laborais individuais*, 4.ª ed., Coimbra.

- ROMANO MARTINEZ, P. (2015), *Direito do Trabalho*, 7.^a ed., Coimbra.
- SALAZAR CASANOVA, J. F. (2004), “Provas ilícitas em processo civil. Sobre a admissibilidade e valoração de meios de prova obtidos pelos particulares”, *Direito e Justiça*, I, pp. 93 ss.
- SEARLE, J. R. (1969), *Speech Acts*, Cambridge.
- SOTTOMAYOR, M. CLARA (2010), *Invalidez e registo: a protecção do terceiro adquirente de boa fé*, Coimbra.
- WOLF & NEUNER (2012), *Allgemeiner Teil des Bürgerlichen Rechts*, 10.^a ed., München.
- ZITELMANN, E. (1879), *Irrtum und Rechtsgeschäft. Eine psychologisch-juristische Untersuchung*, Leipzig.

We don't serve your kind here: Public Accommodations and the mark of Sodom

Joseph William Singer*

Bussey Professor of Law
Harvard Law School

CONTENTS

Introduction

I. Public Accommodations Law and Private Property

II. Mississippi, Apartheid, and Private Property Law

III. Customer choice and equal protection of the laws

IV. Burbclaves, FOQNES, and King William

V. The mark of Sodom

Conclusion

* Thanks and affection go to Martha Minow, Mira Singer, Greg Alexander, Bethany Berger, Nestor Davidson, Eduardo Peñalver. This project was supported in part by funding provided through the research program at Harvard Law School. © 2014 Joseph William Singer.

Negro citizens, North and South, who saw in the Thirteenth Amendment a promise of freedom—freedom to “go and come at pleasure” and to “buy and sell when they please”—would be left with “a mere paper guarantee” if Congress were powerless to assure that a dollar in the hands of a Negro will purchase the same thing as a dollar in the hands of a white man.¹

—Justice Potter Stewart
Jones v. Alfred Mayer Co. (1968)

Those who say “What’s mine is mine, and what’s yours is yours”; this is the average [type of person], though some say this is the type predominant in Sodom.²

—*Pirke Avot (Ethics of the Fathers)* 5:13

Introduction

Two robots walk into a bar. I know, it sounds like a joke. But I am talking about the famous scene in *Star Wars* in which C-3PO and R2-D2 try to follow Luke Skywalker into a tavern. They are looking for a pilot who can take them to the planet Alderaan to fight the evil Emperor. Obi-Wan Kenobi and Chewbacca enter the bar with no problem but when Luke crosses the threshold with the androids, the barkeep stops them with a sneer. “Hey, we don’t serve their kind here,” he says to Luke, refusing to speak directly to the androids.³ “Listen, why don’t you wait out by the speeder,” Luke tells them, “we don’t want any trouble.”⁴

The scene is surreal. Moments before this act of exclusion, we are treated to the sight of many weird and wondrous beings filling the bar. They obviously come from different planets and are designed to intrigue and surprise us by their diversity. The scene is both familiar and strange. The band plays jazz we find familiar but the musicians look like praying mantises on steroids. It is our world and it is not our world. The camera moves to show us all kinds of creatures. In the face of the incredible multiplicity of beings of all sizes and shapes, the act of exclusion is at once familiar and shocking. At the same time, the gesture strikes the viewer as palpably absurd. Why admit the fellow who looks like the devil, the

1 *Jones v. Alfred Mayer Co.*, 392 U.S. 409, 443 (1968) (upholding the constitutionality of a statute that mandated that all citizens have the rights of white citizens in private sales of property) (footnotes omitted).

2 *ETHICS OF THE SAGES: PIRKE AVOT—ANNOTATED & EXPLAINED*, 5:13 (Rami Shapiro trans., 2006) [hereinafter *PIRKE AVOT*].

3 *STAR WARS* (Lucasfilm 1977).

4 *Id.*

praying mantises, and the belligerent guy who starts a fight with Luke, but then draw the line at intelligent robots? Why would they, alone among the diverse clientele, be unwelcome? We realize we know nothing about the history and culture of a world that would be so welcoming to creatures that would surprise and frighten us while excluding androids we have come to view as our companions and comrades.

“We don’t serve their kind here.”⁵ This is a simple exercise of property rights. Or is it? Recently one of my students went to a club in Boston with two of his friends. The bouncer at the door would not let them in. “We don’t want your kind here,” he said, or something to that effect. It was *Star Wars* all over again. My student and his friends are Korean and that apparently bugged the bouncer. They were confused and asked him to explain and he said again that they were not wanted there. They asked to see the manager and, amazingly in this day and age, the manager backed up the bouncer. Not only did he not let them in, he used a racial epithet to express his animus toward Asians. It was 2013 and they were excluded from a bar in Boston because of their race.

I. Public Accommodations Law and Private Property

We are here to celebrate the fiftieth anniversary of the Civil Rights Act of 1964, of which Title II was the public accommodations law. Congress passed that law about a month after my tenth birthday. Fifty years is a long time but we are not talking about ancient history here. I recall segregation; I recall the passage of the public accommodations law. And fifty years is apparently not enough to change our understanding of property—at least not completely. In recent months, we have seen much controversy over a baker who did not want to sell a wedding cake to a same-sex couple and a photographer who did not want to take photographs at a same-sex wedding.⁶ They claimed both expressive and religious liberty to justify denying services to paying customers and they also asserted the right to run their own businesses as they saw fit, retaining the right to exclude customers whose identity or lifestyle they found objectionable.⁷

5 *Id.*

6 See *Elane Photography, LLC v. Willock*, 309 P.3d 53, 77 (N.M. 2013); *Craig*, No. CR 2013-0008 (Colo. Civ. Rights Comm’n Dec. 6, 2013), https://www.aclu.org/sites/default/files/assets/initial_decision_case_no_cr_2013-0008.pdf.

7 See *Elane Photography*, 309 P.3d at 60 (“*Elane Photography* explains that it ‘did not want to convey . . . the story of an event celebrating an understanding of marriage that conflicts with [the owners’] beliefs.’”); *Craig*, No. CR 2013-0008 (“Respondents . . . contend that their refusal was based solely upon a deeply held religious conviction that marriage is only between a man and a woman . . . Respondents contend that application of the law to them . . . would violate their rights of free speech and free exercise of religion . . .”).

A similar claim was made by a web-based adoption service that sought to deny participation by same-sex couples.⁸ These businesses were open to the general public but apparently not *all* of the general public; the owners claimed the right to be selective in their choice of customers and they argued that constraints on their choice burdened their expressive and religious freedoms.⁹ As owners of property, they had the right to exclude and to waive that right selectively and in line with their religious values.¹⁰

If such arguments had been accepted in 1964, then we might still have segregated restaurants, hotels, movie theaters, and pools in the South. After all, segregationists claimed Biblical support for their position and the right to exclude unwanted strangers from their property was a core element of the claim. This argument has not only been revived by businesses seeking to deny services to LGBT¹¹ customers but apparently has some traction with five members of the Supreme Court. The recent case *Burwell v. Hobby Lobby, Inc.*¹² allowed a closely held corporation to deny insurance coverage for contraception to its employees because the owners of the corporation held religious beliefs under which such contraception was a form of abortion and that facilitating such insurance made them complicit in murder.¹³

For my purpose here, what matters about the *Hobby Lobby* decision is the way the Justices in the majority conceptualized property rights. Similar to how the spending of money is now the equivalent of speech for First Amendment purposes,¹⁴ the purchase of such health insurance represents an impermissible burden on the religious expression of the corporation and hence its owners.

8 Butler v. Adoption Media, LLC, 486 F. Supp. 2d 1022, 1059 (N.D. Cal. 2007).

9 See *id.* at 1058 (“Defendants claim that compelling them to post plaintiffs’ profiles on their “web publication” . . . would similarly constitute compelled speech”); *Elane Photography*, 309 P.3d at 67 (“Elane Photography’s choice to offer its services to the public is a business decision, not a decision about its freedom of speech.”); *Craig*, No. CR 2013-0008 (“Respondents therefore have no valid claim that barring them from discriminating against same-sex customers violates their right to free exercise of religion.”).

10 For a critique of those arguments, see Domenick Scudera, *How Is Discrimination a Religious Freedom?*, HUFFINGTON POST (Feb. 20, 2014, 6:52 PM), http://www.huffingtonpost.com/domenick-scudera/how-is-discrimination-a-religious-freedom_b_4814976.html.

11 At the risk of “exclusion,” I will use both LGBT and “gay” to refer to persons whose sexual orientation is other than heterosexual or whose gender identity does not fit into the traditional categories of “male” and “female.” There are inevitable trade-offs between inclusive language and elegant, felicitous expression and I confess to sometimes erring on the side of simplicity and elegance.

12 134 S. Ct. 2751 (2014).

13 See *id.* at 2766, 2785 (holding that the Affordable Care Act’s contraceptive mandate violated business owners’ religious beliefs under the Religious Freedom Restoration Act).

14 *McCutcheon v. FEC*, 134 S. Ct. 1434, 1441 (2014) (holding that per capita limits on political donations violated the First Amendment).

The close corporation—no matter how big, no matter how many employees, no matter how much it dominates a local economy—represents the property of the owners and is subject to their control.¹⁵ The Court assumed that owners have a right to decide how to spend their own money and the conditions on which they will—and will not—allow non-owners onto their premises.¹⁶

And yet the Supreme Court was careful to note that its decision in *Hobby Lobby* did not affect anti-discrimination laws, at least laws prohibiting racial discrimination. As the majority held, “The Government has a compelling interest in providing an equal opportunity to participate in the workforce without regard to race, and prohibitions on racial discrimination are precisely tailored to achieve that critical goal.”¹⁷ This holding suggests that property open to the public for the purpose of hiring employees is subject to reasonable regulations designed to ensure equal access to employment without regard to race. In effect, this line in the sand distinguishes the private home (where one can exclude people from one’s dinner party because of their race) from places of employment and, one assumes, places of public accommodation (where one cannot indulge in such discrimination). This means that the vision of property as under the control of the “owner” and subject to the owner’s “sole and despotic dominion,” as William Blackstone put it, *cannot be the model for all property*.¹⁸ Why is that?

The question is not rhetorical. The *Hobby Lobby* decision pointedly says nothing at all about sex, sexual orientation, or disability discrimination. And Senator Rand Paul famously said on the Rachel Maddow Show that civil rights laws are problematic because they impinge on both the property rights and the free speech rights of owners.¹⁹ The Supreme Court apparently agrees that owners of

15 See *Hobby Lobby*, 134 S. Ct. at 2768 (“[P]rotecting the free-exercise rights of corporations like Hobby Lobby . . . protects the religious liberty of the humans who own and control those companies.”).

16 See *id.*

17 *Id.* at 2783 (addressing the dissent’s concern about using religious practice as an excuse to discriminate).

18 2 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND *2 (Univ. Chi. Press, 1st ed. 1979) (1765-1769) (“There is nothing which so generally strikes the imagination, and engages the affections of mankind, as the right of property; or that sole and despotic dominion which one man claims and exercises over the external things of the world, in total exclusion of the right of any other individual in the universe.”). Some scholars have questioned whether Blackstone in fact conceptualized property rights as absolute. See Robert P. Burns, *Blackstone’s Theory of the “Absolute” Rights of Property*, 54 U. CIN. L. REV. 67, 69 (1985); Carol M. Rose, *Canons of Property Talk, or, Blackstone’s Anxiety*, 108 YALE L.J. 601, 603 (1998) (explaining the ways in which this famous quote deviates from Blackstone’s actual description of the laws of England).

19 Joseph William Singer, *The Anti-Apartheid Principle in American Property Law*, 1 ALA. CIV. RTS. & CIV. LIBERTY. L. REV. 91, 107 (2011) (stating that Senator Paul was soon forced to retreat from this assertion). Senator Paul’s comment sparked extensive controversy. See Jeff Jacoby, Editorial, *Tough Stand: Freedom to Be Odious*, BOS. GLOBE, May 25, 2010, at 15 (defending Senator Paul’s comment on libertarian grounds); Adam Nagourney & Carl Hulse, *Tea Party Pick Causes Uproar on Civil Rights*, N.Y. TIMES, May 21, 2010, at A1; see also Samuel R. Bagenstos, *The Unrelenting Libertarian Challenge to Public Accommodations Law*, 66 STAN. L. REV. 1205, 1221-22 (2014) (criticizing Senator Paul’s conception of property rights).

businesses can express their religious values and exercise constitutionally protected speech through the ways in which they run their businesses and spend their money, and that anyone who wants to work for them has to follow their rules, albeit with some exceptions.²⁰ You live in my house; you follow my rules. But access to the marketplace without regard to race seems to be an important, indeed crucial, exception to this principle. It is of the highest importance to understand why that is so and what it means for our system of private property. To do that, we should consider what the world would look like if that principle were not true.

II. Mississippi, Apartheid, and Private Property Law

To this day, Mississippi has a statute that gives all businesses the right to choose their customers at will. At the risk of being pedantic, it is worth reading the statute in full.

(1) Every person, firm or corporation engaged in any public business, trade or profession of any kind whatsoever in the State of Mississippi, including, but not restricted to, hotels, motels, tourist courts, lodging houses, restaurants, dining room or lunch counters, barber shops, beauty parlors, theatres, moving picture shows, or other places of entertainment and amusement, including public parks and swimming pools, stores of any kind wherein merchandise is offered for sale, is hereby authorized and empowered to choose or select the person or persons he or it desires to do business with, and is further authorized and empowered to refuse to sell to, wait upon or serve any person that the owner, manager or employee of such public place of business does not desire to sell to, wait upon or serve. The provisions of this section shall not apply to corporations or associations engaged in the business of selling electricity, natural gas, or water to the general public, or furnishing telephone service to the public.

(2) Any public place of business may, if it so desires, display a sign posted in said place of business serving notice upon the general public that “the management reserves the right to refuse to sell to, wait upon or serve any person,” however, the display of such a sign shall not be a prerequisite to exercising the authority conferred by this section.

(3) Any person who enters a public place of business in this state, or upon the premises thereof, and is requested or ordered to leave therefrom by the owner, manager or any employee thereof, and after having been so request-

20 See *supra* notes 12-16 and accompanying text (analyzing the *Hobby Lobby* decision).

ed or ordered to leave, refuses so to do, shall be guilty of a trespass and upon conviction therefor shall be fined not more than five hundred dollars (\$500.00) or imprisoned in jail not more than six (6) months, or both such fine and imprisonment.²¹

The statute was passed in 1956, almost two years after the decision in *Brown v. Board of Education*,²² which was decided six days before I was born. The statute is still on the books and still in effect, at least to the extent it has not been preempted by federal statutes or rendered unconstitutional, a question I shall return to in a moment.

What would the world be like if this law were universal? The answer is that it depends on how widespread discrimination is and who owns places of public accommodation and who does not. Let us suppose discriminatory attitudes are widespread and land is owned primarily by white persons who harbor such attitudes. We know what such worlds are like because we have experienced them; they are the South before 1964 and South Africa at the time its anti-apartheid constitution came into effect in 1997, a scant eighteen years ago.

Professor Randall Kennedy vividly depicts that world in his article *The Civil Rights Act's Unsung Victory*.²³ To get ready to travel to South Carolina, Kennedy's parents packed an elaborate picnic to take in the car. They did so not for celebratory reasons or personal enjoyment, but because it was not clear that they would be able to find a place to eat when they got hungry. There might be restaurants along the way but many would not let them in or would serve them only at the back window or along with discourtesy, disparaging speech, and name-calling. There might not be hotels that would take them or gas stations that would fuel their cars. To be prepared and armed, they would take with them *The Negro Motorist Green Book*, a publication by Victor H. Green that identified establishments that served African American customers.²⁴

In other words, they had to prepare to enter hostile territory where they were often unwelcome. More than that, they were disabled from accessing services that human beings need to live, including, food, shelter, bathrooms, and fuel. Their ability to travel was impaired, as was their ability to eat and sleep. Nor could they pass the day without encountering humiliating, degrading, and abasing interactions that conveyed the message, over and over, that your kind are not

21 MISS. CODE ANN. § 97-23-17 (2011).

22 347 U.S. 483, 493 (1954) ("We conclude that in the field of public education the doctrine of 'separate but equal' has no place."); Act of Feb. 21, 1956, ch. 257, 1956 Miss. Laws 307.

23 Randall Kennedy, *The Civil Rights Act's Unsung Victory*, HARPER'S MAG., June 2014, at 35, 35.

24 See Kent Greenfield, *Hobby Lobby and the Return of the "Negro Travelers' Green Book,"* AM. PROSPECT (Mar. 26, 2014), <https://prospect.org/article/hobby-lobby-and-return-negro-travelers-green-book>, archived at <http://perma.cc/LE6K-JZ2N>.

welcome here.²⁵ At the same time, facilities were available of a sort in the South and a family armed with the *Green Book* stood a chance of finding a private home or hotel where they could stay. What would have happened if such facilities were not available? What if almost all the land was owned by white persons who thought the races should be separate and apart?

That was the case in South Africa at the time of the adoption of the interim constitution in 1993 and was still the case in 1997 when the permanent constitution came into effect. Roughly ten percent of the people owned ninety percent of the land and the division was a racial one that had been created and imposed by law.²⁶ Imagine what things were like in 1997 in South Africa. The new constitution had abolished *apartheid*.²⁷ Gone were the laws that required segregation. Did that have the effect of destroying the *apartheid* system? The answer is no.

At that point, the law of private property was neutral; it did not limit access to land based on race nor did it require owners to make racial distinctions. But the *absence* of a law requiring exclusion based on race does not mean that *apartheid* was gone. If white persons own ninety percent of the land and continue to deny black persons access to their property, then *apartheid* would continue unabated *through the seemingly neutral operation of private property law*. If owners have the right to choose their customers and if they have the right to exclude anyone they wish from their property, then *apartheid* could have continued almost unabated through protecting the right of owners to exclude non-owners from their land. Employers could still have refused to hire people based on their race; public accommodations could have refused to serve people because of their race; real estate owners could have refused to sell or rent to buyers or tenants because of their race. Private property law would have been an effective tool to maintain a society based on racial caste.²⁸

That is the reason that the Mississippi statute is so appallingly out of place in 2015. It is based on a misplaced assumption about the nature of private property in a free and democratic society. It has been a long time since the Thirteenth Amendment was adopted and we have abolished not only slavery but also its

25 See Domenick Scudera, *This Is What Discrimination Feels Like*, HUFFINGTON POST (Feb. 27, 2014, 10:43 AM), http://www.huffingtonpost.com/domenick-scudera/this-is-what-discrimination-feels-like_b_4858240.html, archived at <http://perma.cc/6ZBM-CV7P>; see also Patricia Williams, *Spirit-Murdering the Messenger: The Discourse of Fingerpointing as the Law's Response to Racism*, 42 U. MIAMI L. REV. 127, 129 (1987) (discussing the harms of racially discriminatory service in stores, the rhetoric of privatization as a response to racial issues, and racism as a crime).

26 Q&A: *Land Reform in South Africa*, PBS (July 6, 2010), http://www.pbs.org/pov/promisedland/land_reform.php, archived at <http://perma.cc/Q48C-L9YT>.

27 S. AFR. CONST. First Amendment Act of 1997 (establishing a bill of rights for all South African citizens).

28 On the ways that private property rules embody public law values (including equality), see Gregory S. Alexander, *Property's Ends: The Publicness of Private Law Values*, 99 IOWA L. REV. 1257, 1260 (2014).

“badges and incidents.”²⁹ It has been a long time since the Civil Rights Act of 1866 granted all persons the same right to contract as is enjoyed by white persons and granted all citizens the same right to purchase real and personal property as is enjoyed by white persons.³⁰ Those statutes have been interpreted to require places of public accommodation to grant entrance to customers without regard to race.³¹

On the other hand, most federal courts have interpreted the Civil Rights Act of 1866 not to require storeowners to treat customers equally without regard to race, finding no remedy when stores have engaged in racially discriminatory surveillance of customers, treated them to denigrating remarks, or denied them service.³² And the Civil Rights Act of 1964—the law we celebrate today—appears to provide no help here. It regulates inns, restaurants, gas stations, and places of entertainment but seems not to cover retail stores.³³ If no federal statute preempts Mississippi law, then it remains legal in the state of Mississippi to discriminate on the basis of race in treatment of customers in stores—that is, unless the Mississippi statute is unconstitutional. But how can a law that *fails to regulate owners* constitute an unconstitutional deprivation of equal protection of law? Where is the “state action”?³⁴ The state *requires* nothing; rather, it *empowers* owners to control their own property and it *liberates* them from regulation. Any discriminatory decisions are made by the owners, not the state. Even *Shelley v. Kraemer*³⁵ acknowledged that voluntary private acts of discrimination fall outside the Equal Protection Clause.³⁶

Perhaps we need not worry. Perhaps in this day and age, we can expect almost all businesses to eschew such discriminatory conduct and treatment. Norms have changed and it is simply bad business to treat customers with disdain. And even if a few bad apples hold onto such appalling treatment, customers can still boycott and shop elsewhere. The market will discipline such stores

29 *Jones v. Alfred H. Mayer Co.*, 392 U.S. 409, 439 (1968) (“For [the Thirteenth Amendment’s Enabling Clause] clothed ‘Congress with power to pass all laws necessary and proper for abolishing all badges and incidents of slavery in the United States.’” (quoting *The Civil Rights Cases*, 109 U.S. 3, 20 (1883))).

30 42 U.S.C. §§ 1981-82 (2012).

31 Singer, *supra* note 19, at 94.

32 *Id.* at 94-96 (arguing that courts have an inappropriately narrow conception of the right to contract).

33 42 U.S.C. § 2000a (mandating equal access to “lodgings; facilities principally engaged in selling food for consumption on the premises; gasoline stations; places of exhibition or entertainment; other covered establishments”).

34 *The Civil Rights Cases*, 109 U.S. at 12-14 (“[T]he prohibitions of the [Fourteenth Amendment] are against State laws and acts done under State authority.”).

35 334 U.S. 1 (1948).

36 *Id.* at 13 (“So long as the purposes of those agreements are effectuated by voluntary adherence to their terms, it would appear clear that there has been no action by the State and the provisions of the Amendment have not been violated.”).

and the Internet will go wild in publicizing the conduct. Laws are hardly needed to control such behavior (one might argue). And those who persist will constitute a minority with little ability to make life difficult for African Americans. There is no need, one might think, for a new *Green Book*.³⁷

There are two problems with this argument. First, the idea that the market controls all invidious discrimination is, and has always been, demonstrably false.³⁸ The market responds to the attitudes of customers. If one were interested in profits, *and nothing else*, then the Mississippi law makes the existence of discrimination depend on the amount of discriminatory attitudes that exist in a given community. Such attitudes have waxed and waned with time. It is not as if there is no prejudice (say against Latinos) today. Discriminatory attitudes have persisted. Even today, few African Americans have lived without the experience of being accosted by the police, followed by security officers in a store, or treated badly from time to time by others who hold demeaning assumptions about them.³⁹ Studies show that unconscious discrimination continues to haunt us and that job applicants with equal resumes face different prospects because of assumptions about the race of the applicant.⁴⁰ Even my students and colleagues at Harvard Law School and alumni of Harvard University have faced such treatment.⁴¹

Of course this means that neither law nor markets are sufficient to eradicate invidious discrimination. We live, after all, in an age when discriminatory treatment is illegal in most of the country, including the overwhelming majority of states that prohibit discriminatory conduct of any kind in retail stores.⁴² That

37 See *Brooks v. Chi. Downs Ass'n*, 791 F.2d 512, 519 (7th Cir. 1986) (finding that a place of entertainment had no obligation to serve the public because “the market here is not so demonstrably imperfect that there is a monopoly or any allegation of consumer fraud”); Richard A. Epstein, *Public Accommodations Under the Civil Rights Act of 1964: Why Freedom of Association Counts as a Human Right*, 66 STAN. L. REV. 1241, 1290 (2014) (arguing for allowing “private groups to select their own members and govern their own organizations when they provided uniquely differentiated services in competitive markets”).

38 See, e.g., Shalia Dewan, *Discrimination in Housing Against Nonwhites Persists Quietly, U.S. Study Finds*, N.Y. TIMES, June 12, 2013, at B3 (discussing subtle discrimination faced by nonwhite homebuyers and renters); Catherine Ruetschlin, *Markets Don’t Stop Racism but They Can Perpetuate It*, DEMOS (Aug. 21, 2014), <http://www.demos.org/blog/8/21/14/markets-dont-stop-racism-they-can-perpetuate-it>, archived at <http://perma.cc/DZ42-FHDU> (“The classically oblivious treatment of racism as an inefficiency that will be eroded in markets by the pursuit of profits has failed to produce results.”).

39 See Anne-Marie G. Harris, *Shopping While Black: Applying 42 U.S.C. § 1981 to Cases of Consumer Racial Profiling*, 23 B.C. THIRD WORLD L.J. 1, 1 (2003) (discussing the practice of consumer racial profiling and advocating the use of 42 U.S.C. § 1981 to combat it).

40 Audrey J. Lee, Note, *Unconscious Bias Theory in Employment Discrimination Litigation*, 40 HARV. C.R.-C.L. L. REV. 481, 483-86 (2005) (examining the psychology of unconscious bias and its application in litigation strategy).

41 See, e.g., Abby Goodnough, *Harvard Professor Jailed; Officer Is Accused of Bias*, N.Y. TIMES, July 20, 2009, at A13 (addressing the Professor Henry Louis Gates arrest controversy).

42 See Joseph William Singer, *No Right to Exclude: Public Accommodations and Private Property*, 90 NW. U. L. REV. 1283, 1437-39 (1996) (reviewing state statutes that prohibit discriminatory conduct in retail stores).

brings us to the second reason it is wrong to believe that we have no need for laws that prohibit discriminatory conduct in this day and age. The idea that one can “just go elsewhere” *misses the point entirely*. The question is not whether one can find a store willing to let you in and treat you with dignity.⁴³ *The question is whether one has a right to enter stores without worrying about such things.*⁴⁴ The idea that storeowners are the lords of their castles—empowered to control their territory, exclude whoever they wish, and treat those inside as they deem proper and appropriate—contradicts the idea that every person is entitled to the same right as is enjoyed by white persons to enter the *public world of the market* without being treated as a being who is not human or is a member of a lower caste.⁴⁵

Toward the end of the nineteenth century, a myth developed that the public accommodation duty to serve the public was based on the need to respond to monopolies in areas of transportation and inns where no alternatives might be available, yet the need for service was widespread.⁴⁶ That idea was used to justify narrowing the definition of a public accommodation to inns and common carriers; and, *not coincidentally*, this new conception allowed and encouraged new laws that allowed or required racial discrimination in places of entertainment, retail stores, and restaurants.⁴⁷ But that interpretation of public accommodation law as combatting monopoly does not accord with its historic origins, which were based on the moral obligation of businesses that were open to the public to serve the public without discrimination.⁴⁸ The monopoly theory suggests that, if we have competition, then “the market will take care of it” and new businesses will emerge that will serve the excluded group.⁴⁹ It assumes that such groups should be content to go to places where they are welcome. This theory is inconsistent with the actual reason for public accommodation laws. These laws are not about giving despised groups a market niche; they are designed to ensure access to the world of the market without regard to invidious discrimination.⁵⁰

43 *Contra* Epstein, *supra* note 37, at 1241, 1284 (arguing that public accommodation laws should combat monopolies but not interfere with discriminatory businesses, as long as competition leads to other businesses that do not discriminate—at least when segregationist impulses are not common).

44 *See, e.g.*, Singer, *supra* note 19, at 108.

45 *Id.* at 108-09 (arguing for a “legal baseline that rejects market relations premised on unequal status”).

46 *See* Singer, *supra* note 42, at 1390, 1404-06 (“[B]efore the Civil War, the rule that common callings—at least innkeepers and common carriers—had a duty to serve the public was universal . . .”).

47 *See id.* at 1391-95, 1402-03.

48 *See generally id.* (detailing the nineteenth-century history of the public accommodation doctrine).

49 *See* Epstein, *supra* note 37, at 1241.

50 *See* Singer, *supra* note 42, at 1300-01 (arguing that owners have no general right to exclude people unreasonably if they open their property to the general public).

Nor do public accommodation laws constitute invasive interferences with individual freedom or private property rights. Freedom of association is a bedrock constitutional principle but it is not a basis for claiming a right to establish market structures that are premised on invidious discrimination.⁵¹ Our constitutional structure distinguishes between areas of social and political life where groups are presumptively entitled to be exclusionary (such as religion or political associations) and areas of life where access without regard to race or other caste designations is presumptively prohibited—and the main area of life to which the equal access norm applies is the parts of the economy that are open to the general public.⁵²

We have difficult line-drawing problems, to be sure. A wedding band may want to specialize in Klezmer music for use at Jewish weddings. A bakery may seek to sell goods only for use in religious services and seek to limit its wares to Christians. Because the First Amendment protects religious association, businesses that serve *solely* religious purposes may be entitled to do so. But businesses that generally offer their services to the public for all types of service, religious or nonreligious, enter a different area of social life. A bakery that sells goods to customers of all types for all purposes may well be obligated to serve the public regardless of the religion or sexual orientation of the customer.⁵³

This is not to say that legislatures may not make reasonable distinctions between businesses that are obligated to serve the public without discrimination from those that may be allowed to define their customer base more exactly based on religious practice or doctrine. It does mean that we must define the parts of social life where we allow owners to exclude based on religion, political views and affiliation, or other associational interests, from the parts of social life where we require owners to open their property to anyone who is willing and able to seek their services in a non-disruptive manner. One thing is clear: stores that sell their wares to the public, which are not religious establishments, are firmly on the public side of the line.⁵⁴

51 Epstein, for example, is wrong to suggest that public accommodation laws exceed their legitimate scope when they interpret “freedom of association” in an exclusionary manner in the marketplace. See Epstein, *supra* note 37, at 1247.

52 See Singer, *supra* note 42, at 1421 (“The purpose of the [Civil Rights Act] is to afford equal access to businesses that serve the general public.”).

53 See Craig, No. CR 2013-0008 (Colo. Civ. Rights Comm’n Dec. 6, 2013), https://www.aclu.org/sites/default/files/assets/initial_decision_case_no._cr_2013-0008.pdf (finding that a bakery is a place of public accommodation and required to serve the plaintiffs).

54 See, e.g., *id.*

We live in a free and democratic society and that means that certain types of social relationships have been abolished. We no longer have slavery, feudalism, or dictatorial paternal control of familial property and family members (such as wives and children); we have abolished the fee tail, granted women rights in property acquired during marriage, and prohibited indentured servitude, titles of nobility, caste systems, and monopolistic control of land.⁵⁵ Laws outlawing these arrangements can only be effective if we also shape the law of private property to ensure that it cannot be a means to effectuate these banished social arrangements.⁵⁶

What is demanded by a free and democratic society is the right not to experience the humiliation of being turned away from a place open to all others because of characteristics about oneself that should be irrelevant to the opportunity to buy a shirt in a store. This does not amount to “forced association.”⁵⁷ That framing of the issue suggests that storeowners are free in a democracy to choose their customers at will. But that is not the case. Allowing stores to choose their customers at will deprives excluded groups from the freedom to walk into a store that appears to be open to the public and get service.⁵⁸ If “forced association” is at issue, then consider that the absence of a public accommodation law forces a patron to look for some other place that will let her in.⁵⁹ Freedom is not just negative, or the freedom of property owners from regulation of use of their property; freedom is also positive and includes the freedom to enter the marketplace on the same terms as those who do not have to worry about arbitrary exclusion because of the color of their skin.⁶⁰

Telling someone they can just go to another store or try another nightclub is like telling someone to eat soup with a fork. It is a solution that does not solve the problem; it is an answer to the wrong question. The issue is not whether the customer is likely to find another store that will take her in. The question is whether a storeowner has a right, in a free and democratic society, to treat a customer like a pariah. The answer is no.

55 See, e.g., U.S. CONST. art. I, § 9, cl. 8 (banning titles of nobility); U.S. CONST. amend. XIII (abolishing slavery); Singer, *supra* note 19, at 106-07 (arguing that these reforms established a legal baseline for a free society).

56 Singer, *supra* note 19, at 107-08.

57 Epstein, *supra* note 37, at 1256 (objecting to such laws because they lead to “negative-sum games”).

58 See Singer, *supra* note 19, at 109.

59 See *id.*

60 See *id.* (“It should be abundantly evident that the basic policy of United States law is to grant equal access to the marketplace without regard to race.”); Ian Carter, *Positive and Negative Liberty*, STAN. ENCYCLOPEDIA PHIL. (Mar. 5, 2012), <http://plato.stanford.edu/entries/liberty-positive-negative/> (explaining positive freedoms).

III. Customer choice and equal protection of the laws

What does this mean for the Mississippi statute? It means that the statute harkens back to a way of life we have abandoned for good. The statute embraces values we have rejected and a conception of the marketplace and private property that is, in principle, incompatible with a society that has formally abolished the “badges and incidents of slavery.”⁶¹ But now the vaunted state action doctrine rears its ugly head. When the state *authorizes* choice, how is it *denying* rights? The answer has already been given. South Africa would have never gotten rid of *apartheid* if it had left private property owners free to discriminate on the basis of race in public accommodations, employment, and housing. Allowing owners to make choices when they control property that is rightly in the public sphere delegates sovereign power that cannot be delegated.⁶² It turns owners into lords and our Constitution prohibits any government to grant any title of nobility.⁶³

Aficionados of “original intent” theories of constitutional interpretation fail to understand the vast extent to which our norms of equality have changed over time. The Supreme Court has struck down laws that give husbands the power to control their wives’ property—laws that were commonplace both at the time the Equal Protection Clause was adopted in 1791 and when the Fourteenth Amendment came into force in 1868.⁶⁴ It has rejected the “separate but equal” doctrine⁶⁵ and segregationist zoning laws.⁶⁶ It has rejected discriminatory but facially neutral laws such as state laws that enforced racially restrictive covenants and prohibitions on marriage between persons of different races.⁶⁷ It has prohibited discrimination based on disability.⁶⁸

61 The Civil Rights Cases, 109 U.S. 3, 20 (1883).

62 See Morris R. Cohen, *Property and Sovereignty*, 13 CORNELL L.Q. 8, 10-11 (1927) (arguing that laissez-faire judicial doctrine prohibiting minimum wage laws on property rights grounds passed “a certain domain of sovereignty from the state to the private employer of labor”).

63 U.S. CONST. art. I, § 9, cl. 8; see also Joseph William Singer, *Titles of Nobility: Poverty, Immigration, and Property in a Free and Democratic Society*, 1 J.L. PROP. & SOC’Y 1, 13 (2014) (analogizing the modern treatment of immigrants to medieval nobility’s treatment of commoners).

64 See Singer, *supra* note 42, at 1299.

65 *Brown v. Bd. of Educ.*, 347 U.S. 483, 495 (1954) (“We have now announced that such segregation is a denial of the equal protection of the laws.”).

66 *Buchanan v. Warley*, 245 U.S. 60, 82 (1917) (“We think this attempt to prevent the alienation of the property in question to a person of color was not a legitimate exercise of the police power of the State, and is in direct violation of the fundamental law enacted in the Fourteenth Amendment . . .”).

67 *Loving v. Virginia*, 388 U.S. 1, 2 (1967) (holding that anti-miscegenation laws violate the Equal Protection and Due Process Clauses of the Fourteenth Amendment); *Shelley v. Kraemer*, 334 U.S. 1, 20 (1948) (holding that judicial enforcement of racially restrictive covenants violates the Equal Protection Clause of the Fourteenth Amendment).

68 *City of Cleburne, Tex. v. Cleburne Living Ctr.*, 473 U.S. 432, 450 (1985) (finding that the city denied the

Complementing these historical changes in equal protection law has been expanding statutory protection from invidious discrimination. While the Civil Rights Act of 1964 has a short list of public accommodations, the Americans with Disabilities Act of 1990 has a much longer list that includes retail stores, doctors and lawyers' offices, universities, and insurance companies.⁶⁹ I have previously argued that the 1964 Act is ambiguous as to whether its list is exhaustive or illustrative.⁷⁰ Given the changing statutory definitions of what is a public accommodation, I would advise any judge who confronted the issue to interpret the 1964 act to include retail stores. In my view, the 1964 act has been implicitly amended by the later statutes that clarify Congress's current view about what is and is not a "public accommodation" with obligations to serve the public without invidious discrimination.⁷¹ Failing that, I would hold the Mississippi anti-public-accommodations law to be an unconstitutional denial of equal protection of the laws insofar as it authorizes places open to the public to deny service on the basis of race.⁷²

Wesley Hohfeld taught us that privileges are as much laws as rights are laws.⁷³ Privileges confer liberties and may seem to remove state regulation.⁷⁴ Under that view, law only reaches so far as it coerces us to act in one way or another. That means much of life is unregulated by law and beyond the reach of the Equal Protection Clause. But Hohfeld taught us that liberties are laws insofar as they leave others vulnerable to the effects of the exercise of those

respondent's building permit on the basis of animus against the mentally disabled, and that such a denial violated the Equal Protection Clause of the Fourteenth Amendment).

69 42 U.S.C. § 12181 (2012).

70 See Singer, *supra* note 42, at 1412-24.

71 See *id.*

72 Cf. JOHAN VAN DER WALT, *THE HORIZONTAL EFFECT REVOLUTION AND THE QUESTION OF SOVEREIGNTY* (2014) (exploring recent developments in constitutional jurisprudence in various nations that incorporate regulation of private relationships); Joseph Fishkin & William E. Forbath, *The Anti-Oligarchy Constitution*, 94 B.U. L. REV. 669, 670-71 (2014) (arguing that constitutional equality norms have consequences for laws that create or reinforce class inequalities and oligarchic concentrations of power or that prevent individuals from having an equal chance to acquire wealth and participate in economic life); Stephen Gardbaum, *The "Horizontal Effect" of Constitutional Rights*, 102 MICH. L. REV. 387, 390-91 (2003) (arguing that the "law" part of "equal protection of law" includes common law and that includes the law of private property and other laws that regulate relations among persons); Stephen Gardbaum, *Where the (State) Action Is*, 4 INTL. J. CONST. L. 760, 762 (2006) (discussing various conceptions of the application of constitutional norms to relations among private parties and laws regulating those relationships); Paul Gowder, *Equal Law in an Unequal World*, 99 IOWA L. REV. 1021, 1023-24 (2014) (arguing that rule of law ideals require attention to questions of substantive equality in social relationships).

73 See Wesley Newcomb Hohfeld, *Some Fundamental Legal Conceptions as Applied in Judicial Reasoning*, 23 YALE L.J. 16, 30-37 (1913) (discussing the relationship between privileges and rights); Joseph William Singer, *The Legal Rights Debate in Analytical Jurisprudence from Bentham to Hohfeld*, 1982 WIS. L. REV. 975, 986 (analyzing Hohfeldian theories of rights).

74 See Hohfeld, *supra* note 73, at 30-37 ("[A] privilege is the opposite of a duty.").

liberties.⁷⁵ Recall that South Africa could and would have continued apartheid if private property owners were free to exclude others at will. The Mississippi statute denies equality because it denies persons the “same right to contract” and to “purchase real and personal property” as persons who do not face systemic discrimination. Those with the right to exclude exert power over non-owners and this exercise of a right to exclude is not a self-regarding act. Done in concert with others’ acts, it creates a racial caste system.

The Mississippi statute was designed to promote a racial caste system. It enshrines a right that is incompatible with the law of a free and democratic society that treats each person with equal concern and respect and which rejects racial caste, *apartheid*, and segregation. The Mississippi statute, if it means what it says, would require someone to write and to publish an updated *Green Book*. As things stand, LGBT persons actually need such books. About half the states allow discrimination based on sexual orientation and no federal law stands in the way of such discrimination.⁷⁶ Many websites describe places where gay people can feel welcome and implicitly or explicitly let people know where they are not wanted.⁷⁷ We have not moved entirely beyond the world where the *Green Book* was needed.

We spend so much time lauding freedom of choice that we sometimes forget that “[w]e hold these truths to be self-evident, that all [human beings] are created equal” and that a society that has abolished lords and commoners and racial caste has made a decision of consequence to daily life.⁷⁸ Our fundamental commitment to the equal status of all persons means that some property rights cannot be created in a free and democratic society. And the Mississippi statute therefore is not a “deregulatory” law that grants owners freedom to act as they like. It is a regulatory law that *establishes a private property right that is itself inconsistent with equal protection of law*. The Mississippi statute does not deregulate or liberate; it empowers individuals to create property rights that entail domination over others. It authorizes oppression and that cannot withstand equal protection analysis, as I understand it.

Giving owners the freedom to treat customers as they wish and to exclude on any basis they choose is incompatible with the minimum standards for social and civil relationships characteristic of free and democratic societies that have abolished racial castes. Such societies do not allow apartheid to operate

75 See Singer, *supra* note 73, at 987-89 (“[I]f A has the privilege to do certain acts or to refrain from doing those acts, B is vulnerable to the effects of A’s actions.”).

76 JOSEPH WILLIAM SINGER, PROPERTY §2.6.4, at 78 (4th ed. 2014).

77 See, e.g., GAYFRIENDLYGUIDE.COM, <http://www.gayfriendlyguide.com> (last visited Apr. 24, 2015), archived at <http://perma.cc/FWN6-HYJN> (listing businesses that are “genuinely gay-welcoming”).

78 THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776).

through means of private property law. *No law, including private property law, can establish a caste system.* Property that is not open to the public is not subject to equality norms but property that is open to the public becomes subject to the law of civil rights.⁷⁹ The freedom to exclude someone from a public accommodation on the basis of race is simply not a property right that a free and democratic society can recognize, any more than it recognizes titles of nobility or heredity public offices. And any statute that authorizes the creation of such a property right is therefore unconstitutional.

IV. Burbclaves, FOQNES, and King William

Neal Stephenson's *Snow Crash* is one of the most ingenious and influential science fiction novels ever written.⁸⁰ In addition to describing the Internet and Second Life before they even existed, he imagined a world of fragmented sovereignty and absolute property rights.⁸¹ In his fictional world, the United States no longer holds full sovereignty over its territory; rather, sovereignty is divided among many claimants that include corporations, churches, and private armies.⁸² The land is dotted with "Burbclaves" or private, enclosed towns that exercise the right to exclude and rule absolutely within their borders.⁸³ Those towns may ally with each other and give their citizens the freedom to enter their scattered towns and territories. One could be a citizen of McDonald's, for example, or the Catholic Church and be accepted in the Burbclaves that pledge allegiance to those sovereigns.

Stephenson also invented the idea of the franchise-owned quasi-national entities ("FOQNES").⁸⁴ These corporate entities effectively declared independence from state and federal law and became sovereigns of their own. Inside its territory, the law of McDonald's rules supreme, immune from external regula-

79 See Singer, *supra* note 19, at 93 (positing that our conception of the obligations that businesses have changed after the passage of the Civil Rights Acts).

80 NEAL STEPHENSON, *SNOW CRASH* (1992); see also *The Word on Snow Crash and Google Earth*, REALITYPRIME (Sept. 27, 2007, 1:07 AM), <http://www.realityprime.com/blog/2007/09/the-word-on-snow-crash-and-google-earth/>, archived at <http://perma.cc/6FXT-9BA8> (explaining how *Snow Crash* inspired aspects of Google Earth).

81 See STEPHENSON, *supra* note 80, at 24 (describing the "Metaverse," a fictional analogue to the Internet).

82 See *id.* at 44-45.

83 See *id.* at 6 (describing the undesirability of living outside of the Burbclaves); *id.* at 13 (illustrating how security police exclude outsiders); *id.* at 48 (portraying a Burbclave police officer threatening deadly force against a protagonist).

84 *Id.* at 14-15.

tion.⁸⁵ The law of the road was another thing entirely; roads are privately owned and you can choose which road company you want to use.⁸⁶ Once you do, you follow their rules—or their lack of rules.

Why do I bring up Neal Stephenson and the *Snow Crash* world? I do so because the Mississippi statute depicts a world that could devolve into the world of Burbclaves and FOQNEs. That world resembles the feudal state created by William the Conqueror. King William granted lords power over their territory; that power combined what we would consider to be sovereign governing power with the rights that go along with ownership of property. Such lords had the right to exclude others and to rule absolutely inside the manor as long as they complied with their obligations to the king. The same is true of the Burbclaves and FOQNEs. They not only control land; they *rule* it. They determine *who* can enter and *what* they can do there. While that seems innocent when applied to the private home, it leads to pernicious consequences when extended to whole towns or suburbs.

Rather than existing as a citizen of a state who has the freedom to travel, to choose where to live, and to know that one will be treated equally with others wherever one chooses to settle, the *Snow Crash* world is a world of castes. You must petition to become a citizen of a FOQNE or other fragmented state.⁸⁷ *If they let you in*, you have the protection they offer; if they do not let you in, you are out in no man's land or some corporation's road. It is not clear that there is a place where you are entitled to be. And as Jeremy Waldron taught us, "[e]verything that is done has to be done somewhere."⁸⁸ If property law does not ensure access to property *somewhere*, then the law has outlawed your existence. It has made a person illegal. Democracies may regulate conduct but they cannot make it illegal for a person to exist. In the *Snow Crash* world, where you can travel depends on *where you are welcome*. Burbclaves that have many locations allow their citizens relatively open freedom of movement but those that are small or do not have many locations therefore have citizens that cannot leave *without depending on the kindness of strangers*.

What does it mean to never know, when one enters a store, whether one is welcome? How does it affect us if we cannot count on being able to buy food, or clothing, or a computer? How will our life chances and worldview change if our ability to obtain the things we need depended on how much prejudice there was against us? I grew up in the state of New Jersey because my father could not find a company that would hire a Jewish engineer in the early 1950s. But the

85 *Id.* at 44-45.

86 *See id.* at 7.

87 *See id.* at 6-7.

88 Jeremy Waldron, *Homeless and the Issue of Freedom*, 39 UCLA L. REV. 295, 296 (1991).

federal government hired Jews at the Electronics Command at Fort Monmouth, and both Bell Laboratories and Edison Laboratories hired Jewish engineers. My father and mother left New York City because no one would let them in. New York did not prohibit discrimination in employment against Jews. This means it authorized the creation of private property rights that were closed to those of a certain caste. Without a job one cannot eat or live. New York adopted a law that freed its companies to discriminate. That authorized New York to create property rights that gave companies power over others. Because New York adopted this private property system, my father and mother could not stay there. *New York evicted my parents.*

We live in a world that still retains a high degree of discrimination against gay people. And almost half the states have no laws prohibiting discrimination on the basis of sexual orientation; nor do federal laws provide such protection.⁸⁹ Some gay people can avoid discrimination by covering or staying in the closet. They can “pass” and enter commerce in secrecy. Of course, that kind of self-suppression has a cost.⁹⁰ And in those states that do not have antidiscrimination laws, LGBT persons who cannot pass, or do not want to do so, never know as they walk down the street what stores will welcome them or who will become violent if they hold hands or express affection in public. The world is uncertain, the welcome mat invisible, and life often nasty and brutish. Safe places are becoming more common. Anyone who enters a place like Provincetown, Massachusetts, can experience what it is like to walk into a free state where people do not have to hide and where they do not have to fear exclusion or rejection from the properties we need to enter during our daily lives. Public accommodation laws, including the 1964 Civil Rights Act, intend to create such a world; they intend to destroy or prevent the establishment of Neal Stephenson’s fragmented reality and partial citizenship. One cannot live freely, safely, or comfortably if one needs a *Green Book* to travel through one’s day.

89 *Employment Non-Discrimination Act*, HUM. RTS. CAMPAIGN (Dec. 18, 2014), <http://www.hrc.org/resources/entry/employment-non-discrimination-act>, archived at <http://perma.cc/9TE7-LS9R> (“There is no federal law that consistently protects LGBT individuals from employment discrimination; there are no state laws in 29 states that explicitly prohibit discrimination based on sexual orientation . . .”).

90 Brian Krans, *Coming Out of the Closet Could Help You Live Longer*, HEALTHLINE (Feb. 4, 2013), <http://www.healthline.com/health-news/coming-out-increases-mental-well-being-020413>, archived at <http://perma.cc/XL9B-R4AK> (stating that people who are open about their sexuality “experience fewer symptoms of depression, anxiety, and burnout”).

V. The mark of Sodom

The only time the Supreme Court considered the question of whether public accommodation laws violate the U.S. Constitution because they constitute takings of property without just compensation was in the case of *Heart of Atlanta Motel, Inc. v. United States*.⁹¹ In that case, the Court dismissed the takings claim in a single sentence. The opinion explains simply: “Neither do we find any merit in the claim that the Act is a taking of property without just compensation. The cases are to the contrary.”⁹² The cases cited by the Court concern cases where the need for government regulation was evident and important and the uncompensated burden on the owner was justified by the reasons underlying the public regulation. As to public accommodations laws that prohibit race discrimination, the public interest is apparently compelling, so clearly so that the majority of the Supreme Court in *Hobby Lobby* could dismiss, in another single sentence, the idea that anyone might assert religious interests in racial discrimination in employment.⁹³ “The Government has a compelling interest in providing an equal opportunity to participate in the workforce without regard to race, and prohibitions on racial discrimination are precisely tailored to achieve that critical goal.”⁹⁴

While one might argue that the interest in equal access to public accommodations, employment, and housing justifies limiting property rights, a more appropriate conclusion is that “[p]roperty rights serve human values. They are recognized to that end, and are limited by it.”⁹⁵ Contrary to what Senator Paul believes, civil rights laws do not limit property rights.⁹⁶ They define what property rights can exist in a free and democratic society. They establish the structural baseline, the infrastructure of a society that is committed to granting equal protection of the laws.

In the world of King William and in the world of FOQNEs, property rights are as absolute as the various powers that be can make them. Those who own/rule

91 379 U.S. 241, 261-62 (1964) (upholding a public accommodations law on basis of the Commerce Clause).

92 *Id.* at 261 (citing *United States v. Cent. Eureka Mining Co.*, 357 U.S. 155, 168-69 (1958) (holding that wartime restrictions during World War II on the operation of gold mines did not constitute takings); *Omnia Com. Co. v. United States*, 261 U.S. 502, 510 (1923) (holding that government interruption of a steel production contract for wartime needs during World War I did not constitute a taking); *Legal Tender Cases*, 79 U.S. (12 Wall) 457, 551 (1870) (holding that paper money printing during the Civil War did not constitute a taking)).

93 *Burwell v. Hobby Lobby, Inc.*, 134 S. Ct. 2751, 2783 (2014).

94 *Id.*

95 *State v. Shack*, 277 A.2d 369, 372 (N.J. 1971) (overturning criminal trespass convictions for defendants who walked on to a farmer’s land to give medical and legal aid to indigent farmworkers).

96 See *supra* notes 12-16 and accompanying text.

towns can act as they please with their own—and with those they allow to enter their domains. *But that is not our world.*

In *our world—the world where we hold some truths to be self-evident*—the power to exclude or to treat someone in a disparaging manner because of their race does not extend to business property open to the public. It is not a property right that can be recognized or granted legal recognition in a free and democratic society. It is incompatible with a society that has abolished racial caste and which ensures in its Constitution that all persons are entitled to equal protection of the laws—a right interpreted by the Reconstruction Congress as encompassing the rights to contract and to purchase personal property on the same basis as is enjoyed by those who have not suffered historic discrimination.⁹⁷

A state that pretends to grant everyone equal rights but that in effect denies a segment of the population the liberties, the freedoms, the securities associated with daily life deals in tarnished goods. We might better understand this point by considering a *midrash*, a rabbinic story, about the reason for the destruction of the town of Sodom. The rabbis argued that Sodom was destroyed *because it did not understand the limits of property rights*.⁹⁸

The Talmud says that the person who is strict about property rights says “[w]hat’s mine is mine, and what’s yours is yours.”⁹⁹ But then the Talmud goes on to say something startling: “some say this is the mark of the people of Sodom.”¹⁰⁰ For a religion like Judaism that accepts private property, that makes the injunction not to steal one of its ten most central commands,¹⁰¹ how could respect for “mine and thine” be the mark of a people so irredeemable that God saw fit to destroy them?

Ezekiel explains that Sodom was rich but cruel. “[S]he and her daughters had pride, surfeit of food, and prosperous ease, but did not aid the poor and the needy.”¹⁰² The rabbis tell us the people of Sodom prohibited charity; they took “what is mine is mine” to an extreme. When the people of Sodom saw a young woman give food to a starving neighbor, they burnt her alive.¹⁰³ Charity was against the law in Sodom.¹⁰⁴ What is mine is mine and what is yours is yours.

97 See *supra* note 30 and accompanying text (discussing the effect of the Civil Rights Act of 1866).

98 See PIRKE AVOT, *supra* note 2, at 5:13 (“Those who say ‘What’s mine is mine, and what’s yours is yours’; this is the average [type of person], though some say this is the type predominant in Sodom.”).

99 *Id.*

100 *Id.* (contrasting one who is strict about property rights to the fool, the saint, and the wicked types) (translation by author).

101 Exodus 20:15 (“Thou shall not steal.”).

102 Ezekiel 16:49.

103 MIDRASH RABBAH, Genesis (Vayera), 49:6 (Rabbi Dr. H. Freedman, trans., Soncino Press 3d ed. 1983).

104 *Id.* at 49:6 n.3 (“[T]hey had strictly forbidden charity . . .”).

The rabbis tell us further that the people of Sodom gave coins to the poor but they wrote their names on those coins.¹⁰⁵ When the coins were offered to the storeowners for bread, the shopkeepers would see the names and refuse to accept the marked money.¹⁰⁶ The poor had money in their hands, but no one would take it, no one would sell to them, no one would let them in, and the poor would die in the street.¹⁰⁷ And then the residents would come to take back their money.¹⁰⁸

The Talmud tells us that what is mine is mine and what is yours is yours.¹⁰⁹ But the Torah also tells us: *Al tonu ish et akhiv*. Do not wrong your brother.¹¹⁰ This commandment is the source of many Talmudic regulations mandating fair transactions in the marketplace. And the obligation extends to strangers as well as kin. After all, Leviticus 19:34 tells us that “[t]he stranger who resides with you shall be as one of your citizens; you shall love him as yourself.”¹¹¹ We are obligated to treat others as we would want to be treated. That principle applies as much to the marketplace as to other areas of life, and it applies to strangers as much as it does to citizens.

Our society embraces these norms by requiring all laws to recognize the equal status of each person. The society that provides stores and businesses open to the public cannot contradict that message by allowing owners to exclude people unless they have good reason to do so. The owner who hangs a shingle and offers her services to the public cannot retreat from the promise of open service; to do so is to offer the public marked money. It is to convey the promise of a free and open society and then take the prize away from the despised few. A free and democratic society abolishes titles of nobility, outlaws social castes and racial apartheid, and promotes equal access to the free market.

105 See THE BABYLONIAN TALMUD: SEDER NEZIKIM, SANHEDRIN, 109a.

106 *Id.*

107 *Id.*

108 *Id.*

109 PIRKE AVOT, *supra* note 2, at 5:13.

110 Leviticus 25:14 (“When you sell property to your neighbor, or buy any from your neighbor, you shall not wrong one another.”); JPS HEBREW-ENGLISH TANAKH (Jewish Publication Society 2000).

111 *Id.* at 19:34.

Conclusion

Markets are free not because they are unregulated but because they are open to all. Property is private not because owners can do anything they like on their property but because property law promotes the legitimate interests of persons who have chosen to live together in a free and democratic society that ensures that each person is due equal concern and respect. Private property that is open to the public and which serves the public provides a vehicle for social life as well as economic production and distribution. Because we do not have a racial caste system, public accommodations must be open to all. Public accommodations law is not a nice extra that we can be happy the Congress passed in 1964. Nor does it limit the rights of property owners. It defines what property rights are compatible with the truths that we hold self-evident and our commitment to freedom, equality, and democracy. The public accommodations law of 1964 has become, and will always be, a bulwark of our democratic system of government and our free and democratic way of life. That is why the Mississippi anti-public-accommodation law cannot mean what it says and that is why it cannot serve as a legal basis for discrimination in a retail store.

Sobre a alegada “Superação” do Direito pela análise económica (ilustrada com a análise das medidas da indemnização contratual)

Paulo Mota Pinto

Professor Auxiliar
Faculdade de Direito da Universidade de Coimbra

SUMÁRIO

- 1. Direito ou análise económica?**
- 2. A teoria do “não cumprimento eficiente”**
- 3. Crítica à teoria do “não cumprimento eficiente”**
- 4. Os efeitos dos “remédios” indemnizatórios na responsabilidade contratual**

1. Direito ou análise económica?

Como se sabe, a *law & economics* constitui hoje ainda o movimento teórico de abordagem do Direito mais poderoso na doutrina norte-americana. Tal corrente procura uma impositação e solução dos problemas jurídicos fundamentalmente a partir da “eficiência de alocação” de recursos, submetendo esses problemas e soluções a uma análise económica, de “bem-estar”, de acordo com os cânones da escola do marginalismo – designadamente, a análise do valor marginal das opções de comportamento de agentes sob uma restrição, e em geral a comparação dos custos e benefícios das soluções legislativas. E complementa esta perspetiva com abordagens teóricas como as da teoria dos jogos e da economia da informação, designadamente, para analisar, em diversos pontos, os incentivos resultantes das regras jurídicas ou a forma como estas possibilitam a coordenação da ação.

Assume tal análise um ponto de vista funcionalista, em que o fim social decisivo é, se não a “utilidade total”, a “maximização da riqueza” (assim, Richard Posner) ou de “bem-estar”, procedendo a uma “economização do Direito” auto-limitada a esses fins (entendidos embora com amplitude, mas deixando de fora os “problemas de distribuição”). E postula, não só um individualismo (metódico e antropológico) como necessariamente uma visão do indivíduo dominado por tal racionalidade utilitária, enquanto maximizador racional dos seus fins, em que aquilo que sente que deve fazer apenas poderia relevar no contexto desta maximização, e o Direito aparece como técnica, operador ou instrumento desse fim.

A referida análise foi, desde os anos 70 do século xx, transportada também para o direito dos contratos – não só para a explicação das funções da disciplina contratual (por exemplo, das normas supletivas, “maioritárias” ou “penalizadoras”), como, designadamente, para a análise dos “remédios” contratuais e pré-contratuais (a análise da força e do momento da vinculação contratual), mesmo em condições de “custos de transacção” elevados¹.

1 A noção de “custos de transacção” foi fundamental para a análise económica do Direito e das instituições. Cf. RONALD COASE, “The Nature of the Firm”, *Economica*, 4, 1937, pp. 386, ss. (reimpr. em ID., *The Firm, the Market and the Law*, Chicago, 1988, pp. 33 a 55) referindo-se ao significado, na afetação de fatores de produção no mercado, desses custos, e do seu relevo para a existência de empresas. Em “The Problem of Social Cost”, *Journal of Law & Economics*, n.º 3, 1960, pp. 1 a 44 (= *The Firm...*, cit., pp. 95 a 156), COASE afirmou o depois denominado “Teorema de Coase”, com duas asserções: num mundo perfeito, sem custos de transacção, o mercado gerará uma eficiente alocação ou afetação de recursos; num tal mundo (“coaseano”), a eficiente afetação de recursos é independente da originária atribuição de direitos sobre os recursos existentes pela ordem jurídica. A ineficiente afetação de recursos pela ordem jurídica seria, pois, irrelevante num mundo “coaseano”: “é sempre possível modificar através de transações no mercado a delimitação jurídica inicial de direitos. E, claro, se tais transações no mercado não têm custos, um tal rearranjo de direitos terá sempre lugar se conduzir a um aumento no valor da produção” (“The Problem of social cost”, in *The Firm...*, cit., p. 114).

Ora, é nossa convicção que os problemas fundamentais do direito dos contratos são problemas especificamente jurídico-normativos, e que devem ser tratados, não só de acordo com as referências próprias de cada sistema, como, designadamente, de acordo com a intenção de sentido que inspira a perspetiva jurídica.

Não pretendemos, com isto, de modo algum excluir o préstimo dos resultados de disciplinas auxiliares, designadamente, no apuramento das motivações das partes e dos efeitos de um determinado regime, a considerar necessariamente para a formulação de determinadas soluções. Tal não seria avisado. Aqueles conhecimentos científicos e esta técnica – consideração das regras através do ponto de vista da maximização de uma função – devem ter um papel auxiliar (sobretudo em opções legislativas ou judiciais onde as implicações ético-materiais são de menor importância), para tornar clara a relevância da realidade a que se dirigem, a estrutura factual, tanto desta, como dos efeitos de um determinado regime. Mas não resolvem o problema jurídico-normativo.

Designadamente, pensamos que a substituição da racionalidade jurídica pela análise económica deixa escapar necessariamente os aspetos decisivos², e fá-lo também na análise da medida dos danos (dos “remédios” pré-contratuais e contratuais).

O que não será grave, sempre que os resultados a que se chegue sejam apresentados, e vistos, como meramente auxiliares – como traduzindo só uma perspetiva, externa, de descrição do problema ou (de acordo com um padrão assumido, que não é jurídico, mas antes a eficiência ou a “maximização da riqueza”) de prescrição de soluções. A não ser assim, isto é, se a análise económica pretender, em perspetiva “normativa”, substituir, ou superar o Direito como disciplina autónoma³, cremos, porém, que estaremos perante um sério desvio metodológico, a revelar grave incompreensão do específico sentido do Direito, quando não mesmo uma opção antropológica discutível e, se levada até ao fim, um retrocesso cultural.

2 Cf. um juízo crítico das insuficiências da análise económica, a propósito da fundamentação e dos limites da declaração negocial, já em *Declaração tácita e comportamento concludente no negócio jurídico*, Coimbra, 1995, p. 168, n. 19. Convém recordar, no entanto, que o movimento da *law & economics* é complexo, com várias fases de desenvolvimento e orientações. V., sobre estas, GARY MINDA, *Postmodern Legal Movements – Law and Jurisprudence at Century's End*, New York, New York Un. Pr., 1995, pp. 85 e ss. (distinguindo uma primeira geração de “jovens turcos” e uma segunda, associada a uma *post-Chicago law and economics*), e, entre nós, JÓNATAS M. MACHADO, *Liberdade de expressão. Dimensões constitucionais da esfera pública no sistema social*, Coimbra, Coimbra Ed., 2002, pp. 202 e ss.

3 Cf. RICHARD A. POSNER, “The Decline of Law as an Autonomous Discipline: 1962-1987”, *HLR*, 1987, pp. 761-80, e ID., *Overcoming Law*, Cambridge: Harvard Un. Pr., 1995. Criticamente, cf. ROBERT S. SUMMERS, “Economics and the Autonomy of Law, Legal Analysis and Legal Theory”, in *Legal Analysis and Legal Theory, Rechtstheorie*, Beiheft 10, Dordrecht, Kluwer, 2000, pp. 395-408.

No plano *metodológico*, estaria, desde logo, em causa a diferença entre a intenção de concretizar uma racionalidade prática e valorativa, do Direito, e a aplicação das conclusões de uma disciplina teórica ou da sua técnica aplicativa. Com a pretensão de traduzir totalmente as valorações jurídicas em categorias económicas, a racionalidade jurídica acaba por ser substituída por uma *técnica*, cujo sucesso depende da relação meio/fim, estando este previamente fixado e deixando a atividade do jurista de ser prudencial (substituição da *phronesis* pela *technê*).

Ora, pode, desde logo, perguntar-se por que razão é a *eficiência*, a “*maximização da riqueza*” ou do *bem-estar* o fim decisivo, perante limites ou constrangimentos normativos como os dos direitos fundamentais⁴. Por exemplo, um contrato pelo qual uma pessoa se coloque numa situação de escravidão para se sustentar a si e à sua família não é certamente justificável juridicamente por uma hipotética maior racionalidade económica, mesmo a longo prazo, de uma tal escolha.

Depois – e consideramo-lo particularmente significativo, por mostrar como a análise económica passa ao lado da compreensão, mesmo que só intuitiva ou não autorreflectida, do Direito –, há que rejeitar a exclusão da análise da justiça das soluções reconhecidas pela ordem jurídica nos seus reflexos sobre a posição de cada sujeito, relegadas em autolimitação para o mero plano “*distributivo*”. A equivocidade desta última designação não deve induzir em erro: tais meros “problemas distributivos” não são apenas aqueles em que possa estar em causa a aristotélica *iustitia distributiva*; trata-se, antes, realmente, dos próprios critérios característicos da racionalidade prática, da justiça na relação entre as partes (incluindo a *iustitia comutativa* e corretiva), da “correção” ou “justeza” (*fairness*) da solução, do equilíbrio axiológico da valoração, etc., tudo dimensões implicadas, até ao nível de uma perspectiva pré-reflexiva, no sentido do Direito (consiga este, ou não, “uma resposta certa”), que não são captadas pela redução deste a uma “regulação” ou “condução de comportamentos” (“*Verhaltenssteuerung*”), para mais segundo uma lógica de eficiência ou de “maximização de riqueza”. Isto é evidente até ao nível dos resultados, em termos de equilíbrio das posições dos intervenientes, pressupostos pela aplicação do “teorema de Coase”: sem “custos de transação” os recursos (os direitos sobre

4 Para este e outros tópicos críticos que referimos, cf. A. CASTANHEIRA NEVES, *Apontamentos complementares de teoria do direito*, Coimbra, policop., 1998, pp. 65-6. V. tb. FERNANDO PINTO BRONZE, *Lições de introdução ao Direito*, 2.ª ed., Coimbra, Coimbra Ed., 2006, p. 566, n. 340. Cf. tb. RONALD DWORKIN, “Why Efficiency? A Response to Professors Calabresi and Posner”, *Hofstra LR*, 8, 1980, pp. 563-90, ID., “Is Wealth a Value?”, *JLS*, 9, 1980, pp. 191-226, ID., *Law’s Empire*, Cambr. (Mass.), Bellknap Pr., 1986, pp. 286 e ss. (e, sobre a contraposição entre princípios e utilidade, já ID., *Taking Rights Seriously*, Cambr., Mass, Harvard Un. Pr., 1977, pp. 94 e ss.). Mas cf. tb., mais recentemente, LOUIS KAPLOW/STEVEN SHAVELL, *Fairness versus Welfare*, Cambridge (Mas.), Harvard Univ. Press, 2002.

eles) ficarão onde a sua afetação é mais eficiente, mas é claro que a afetação inicial é decisiva para as posições relativas das partes, assim ganhando, ou tendo de pagar, na negociação que se seguirá. Tal afetação conduz, pois, a resultados inversos em termos “distributivos”.

Passando sobre os problemas da visão antropológica pressuposta – da tentativa de recriar um “homem unidimensional”, ou que reduz as suas facetas apenas a uma dimensão –, concluir-se-á que as objeções contra esta perspetiva fazem presa mais fundo, na própria autonomia do sistema e da racionalidade jurídicos.

A sua absolutização conduz a uma abdicação da responsabilidade prática do jurista, num instrumentalismo que o torna cego para os padrões de referência especificamente jurídicos – não é, pois, Direito, mas uma *alternativa ao Direito*, e, portanto, como tem salientado Castanheira Neves⁵, uma proposta de dissolução de um projeto humano autónomo (e de um produto cultural), assumido até hoje como compromisso civilizacional inconfundível⁶.

Em particular no direito privado, nota-se que tal perspetiva pode não ser adequada, pois este não está confinado, nem confina as partes, a uma perspetiva de eficiência e “antes protege o uso da autonomia privada nos limites que lhe são destinados também na medida em que o titular individual faça dela uma emprego contrário a valorações de mercado”⁷. O direito civil “não é um puro

5 A reflexão sistematizadora das ameaças à autonomia do Direito (e das alternativas a este), para reafirmação e realização do(s) sentido(s) que histórico-culturalmente o Direito tem tido, como verdadeira característica civilizacional e, portanto, possibilidade cultural e humana, tem sido uma proposta recorrente do pensamento jurídico-filosófico (e também na sua projeção metodológica) de ANTÓNIO CASTANHEIRA NEVES: *O direito hoje e com que sentido? O problema actual da autonomia do direito*, Lisboa, Inst. Piaget, 2002 (e pp. 45 e s., sobre a “análise económica do Direito”); ID., “Coordenadas de uma reflexão sobre o problema universal do direito — ou as condições da emergência do direito como direito”, in *Estudos em homenagem à Professora Doutora Isabel de Magalhães Colaço*, vol. II, Coimbra, Almedina, 2002, pp. 837-71; ID., “A crise actual da filosofia do direito no contexto da crise global da Filosofia: tópicos para a possibilidade de uma reflexiva reabilitação”, *Estudos dedicados ao Prof. Doutor Mário Júlio de Almeida Costa*, Lisboa, Universidade Católica, 2002, pp. 147-280 (e tb. *Studia Iuridica*, n.º 72, Coimbra, Coimbra Ed., 2003); e já antes: ID., *Metodologia jurídica. Problemas fundamentais*, Coimbra, *Studia Iuridica*, Coimbra Ed., 1993, pp. 54, ss.; ID., “Método jurídico”, in *Digesta*, vol. II, Coimbra, Coimbra Ed., 1995, pp. 283-336 (325, ss.); ID., *Teoria do direito. Lições proferidas no ano lectivo de 1998/1999*, Coimbra, policop., 1998, pp. 127, ss.; ID., *Apontamentos complementares...*, cit., *passim*; ID., “Entre o ‘Legislador’, a ‘Sociedade’ e o ‘Juiz’ ou entre ‘Sistema’, ‘Função’ e ‘Problema’ – Os modelos actualmente alternativos da realização jurisdicional do Direito”, in *BFD*, 74 (1998), pp. 1-44. E ainda ANTÓNIO CASTANHEIRA NEVES, “O funcionalismo jurídico – Caracterização fundamental e consideração crítica no contexto actual do sentido da juridicidade”, *RLJ*, ano 136.º (2006), pp. 3-31, 66-86.

6 Já a resposta de RICHARD POSNER às objeções jurídico-filosóficas tem, a nosso ver, de considerar-se desconsoladora, consumando-se na recusa dos problemas pela adoção de uma posição “pragmatista”, que pretendeu filiar em OLIVER WENDELL HOLMES, JR.

7 NORBERT HORN, “Zur ökonomischen Rationalität des Privatrechts. Die privatrechtstheoretische Verwertbarkeit der ‘Economic Analysis of Law’”, *AcP*, vol. 176 (1976), pp. 307-333.

direito de eficiência e também não o deveria ser”⁸. Também, no direito privado, a utilização da análise económica como critério decisivo para problemas jurídicos tem, pois, como consequência a perda de autonomia do “ponto de vista jurídico” particularmente constitutivo deste ramo do Direito.

Pode aceitar-se que os dados de disciplinas como a economia são importantes, e mesmo indispensáveis, para aferir a adequação da regulamentação ao seu “domínio normativo”, condicionando o espaço de liberdade do legislador.

Mas além de a racionalidade não poder ser reduzida a um “controlo do sucesso”, a própria avaliação dos resultados do programa normativo, ou da decisão do caso, pressupõe critérios jurídicos aos quais a análise económica é cega, e, portanto, é também aquele controlo que se torna impossível⁹.

O que se revela, além de em “explicações” como a recondução da razão para a “sofisticação” e complexidade da disciplina contratual aos custos de transação e informação e a erros que impedem a eficiência irrestrita da negociação, também no controlo dos resultados de diferentes medidas da indemnização, no critério para a verificação da existência de uma “fundamental equivalência, em contínuo, de muitas soluções que a dogmática cria serem estanques” – por exemplo, a *dissolução da distinção* entre o interesse na confiança (interesse contratual negativo) e o interesse no cumprimento (interesse contratual positivo), tal como é proposta, numa atitude pragmática e flexibilizadora conducente a uma “despreocupação semântica” (passando tais “interesses” a não ser mais do que um “‘ponto focal’, um referente semântico que facilitaria supletivamente a formação de um conteúdo obrigacional entre as partes”), é o reflexo da perspetiva instrumentalista com que a análise económica aborda o problema da indemnização contratual.

Importa, porém, apurar se existe alguma razão, da perspetiva dogmático-jurídica – e não apenas na da realização de um valor externo, de que os “interesses” em causa sejam função –, para distinguir tais valores “discretos”, em lugar de aceitar apenas um *continuum* de medidas de danos (e embora isso não

8 HANS-BERND SCHÄFFER, “Allokationseffizienz als Grundprinzip des Zivilrechts”, in CLAUD OTT/HANS-BERND SCHÄFFER (orgs.), *Allokationseffizienz in der Rechtsordnung. Beiträge zur ökonomischen Analyse des Zivilrechts*, Berlin, Springer, 1989, pp. 1-24 (19).

9 Cf. tb. JÜRGEN HABERMAS, “Missverständnis der Rationalität als Erfolgskontrolle”, in ID., *Theorie und Praxis*, 2.ª ed., Frankfurt. a. M., Suhrkamp, 1967, pp. 231, ss., Josef ESSER, *Vorverständnis und Methodenwahl in der Rechtsfindung*, Frankfurt a. M., Athenäum, 1970, p. 213 (“segundo o nosso entendimento do Direito, a abdicação de uma *ratio legis* mais do que funcional é simultaneamente renúncia a uma *ratio iudicis* responsabilizante, e, até, nem se poderão erigir padrões de eficiência do Direito, também sob pontos de vista económicos, se se renunciar a critérios explicitáveis para aquilo que, para o Direito, significa um ‘sucesso’ ou um ‘insucesso’”). Sobre a possibilidade de a análise económica fornecer em certos casos mais do que uma solução jurídica igualmente eficiente para o mesmo problema, embora com diversas consequências distributivas, cf. exemplos em MICHEL ROSENFELD, *Just Interpretations: Law Between Ethics and Politics*, Berkeley, Univ. of Calif. Pr., 1998, p. 187.

impeça a possibilidade de reduções ou agravamentos, considerando critérios suplementares). E, ainda quando se admita que tais razões existam, tratar-se-á, porém, de um muito diferente reconhecimento da possibilidade de um *continuum*, que não é determinado pela colocação deste em função de valores exteriores ao Direito. Com efeito, a tendência para esbater “contraposições agudas” e reconhecer fatores que intervêm de modo gradativo pode ser também assumida para a realização de valores diversos da eficiência, próprios do regime jurídico de certo instituto. É, em grande medida, o que levou às propostas de utilização de um “sistema móvel” em diversos domínios – embora para a definição dos pressupostos, e não da consequência, estava aqui também em questão a preocupação de “limar arestas” introduzidas por pressupostos rígidos¹⁰.

Não podem, assim, ser aceitas tentativas de resolver as questões jurídicas fundamentais através do recurso direto a perspectivas de outras ciências.

Os dados de disciplinas auxiliares são importantes, mas não parece que exista, ou pelo menos não foi demonstrada até hoje, uma “predeterminação” unívoca da solução dos problemas jurídicos, designadamente da responsabilidade contratual e pré-contratual, a partir dos conhecimentos económicos, ou de outra disciplina – como, aliás, é patenteado pela própria variedade diacrónica¹¹ e sincrónica das soluções, sem que se mostre que ela é acompanhada de alterações das variáveis correspondentes.

10 Cf., designadamente, WALTER WILBURG, *Die Elemente des Schadensrechts*, Marburg a. d. Lahn, Elwert, 1941, ID., *Entwicklung eines beweglichen Systems im Bürgerlichen Recht*, Graz, Kienreich, 1951, e ID., “Zusammenspiel der Kräfte im Aufbau des Schuldrechts”, *AcP*, 163, 1964, pp. 346-79. Em geral, v. FRANZ BYDLINSKI/CHRISTIAN VON BAR (orgs.), *Das bewegliche System im geltenden und künftigen Recht*, Wien-New York, Springer, 1986, *passim*, FRANZ BYDLINSKI, *Privatautonomie und objektive Grundlagen des verpflichtenden Rechtsgeschäftes*, Wien-New York, Springer, 1967, pp. 122 e ss., ID., “A ‘Flexible System’ Approach to Contract Law”, in HERBERT HAUSMANINGER/HELMUT KOZIOL/ALFREDO M. RABELLO/ISRAEL GILEAD (orgs.), *Developments in Austrian and Israeli Private Law*, Springer, New York, 1999, pp. 9-20, ID., “Mistake in Austrian Private Law Viewed in Terms of a ‘Flexible System Approach’”, in H. HAUSMANINGER et alii, ob. cit., pp. 21-47 (ID., “Das österreichische Irrtumsrecht als Ergebnis und Gegenstand beweglichen Systemdenkens”, *Festschrift für Hans Stoll zum 75. Geburtstag*, Tübingen, Mohr Siebeck, 2001, pp. 113-141), FRANK O. FISCHER, “Das ‘bewegliche System’ als Ausweg aus der ‘dogmatischen Krise’ in der Rechtspraxis”, *AcP*, vol. 197 (1997), pp. 589-608, e FRANZ M. ADAMOVIC, “Das bewegliche System in der Rechtsprechung”, *JBl*, vol. 124, 2002, pp. 681-702.

11 As diferentes interpretações sobre a influência de fatores económicos na evolução do direito dos contratos do próprio *common law* podem ser ilustradas com as divergências entre PATRICK SELIM ATIYAH, *The Rise and Fall of Freedom of Contract*, Oxford, Clarendon Press, 1979, pp. 398, ss., e, para o direito americano, MORTON J. HORWITZ, *The Transformation of American Law, 1780-1869*, cit., pp. 173, ss., e já antes ID., “The Historical Foundations of Contract Law”, *HLLR*, 87, 1974, pp. 917-956, por um lado – defendendo o papel decisivo, na origem da ação contratual, da expansão económica e da pressão do comércio, numa “transformação” do direito dos contratos, com ascensão da teoria da vontade e da liberdade contratual –, e, por outro lado, a tese de uma fundamental continuidade (v. a refutação pormenorizada de ALFRED W. B. SIMPSON, “The Horwitz Thesis and the History of Contracts”, *University of Chicago LR*, 46, 1987, pp. 533-601, e KEVIN M. TEEVEN, *A History of the Anglo-American Common Law of Contract*, New York, Greenwood Press, 1990, pp. 162 e ss.). Cf. ainda, sobre este debate, WYTHE

Aliás, admitida a determinação das soluções jurídicas por ciências “auxiliares”, sempre haveria que perguntar: porque escolher uma e não outra? Porquê recorrer decisivamente à economia, e não, por exemplo, à sociologia ou à psicologia. Sem prejuízo do seu grande relevo (porventura mesmo muitas vezes decisivo), julgamos que não bastaria para tanto a remissão para um entendimento amplo da economia como ciência da racionalidade estratégica, atinente à relação entre fins e meios, que descreve e recomenda o comportamento humano em condições de escassez de meios. Pois sempre haveria que explicar porque se reduz o elemento decisivo a essa relação. A busca de soluções jurídicas apenas em sede extrajudicial, para além de necessariamente redutora, não pode também bastar se com a redução a uma das disciplinas auxiliares da elaboração do programa normativo pelo legislador (ou da decisão judicial).

Os contributos que se procure obter noutras ciências carecem, portanto, de uma “tradução” ao nível jurídico-normativo e dogmático, através de um trabalho que dê conta da perspetiva específica do Direito. Sem um esclarecimento dogmático do modo como se podem fazer frutificar essas contribuições, teremos de dizer, portanto, que a sua repercussão jurídica fica em aberto, e que a tentativa de arvorar aquelas em instância decisiva conduziria à passagem para o plano de ciências auxiliares, descurando o jurídico.

2. A teoria do “não cumprimento eficiente”

Rejeitamos, assim, a perspetiva do instrumentalismo da análise económica, que leva a perder os critérios dogmáticos de uma teoria jurídico-dogmática das diversas medidas da indemnização contratual.

As consequências da análise económica podem, com efeito, ser ilustradas com a sua aplicação ao regime do não cumprimento e das medidas de indemnização¹².

Tratando-se de responsabilidade contratual, é natural que a análise económica se tenha primeiro centrado na otimização da decisão do devedor de *cumprir* ou *não cumprir*, e, portanto, no *cumprimento* e *não cumprimento eficientes*. O “remédio” para o não cumprimento deveria ser escolhido tendo em conta os efeitos que teria sobre as decisões das partes de cumprir ou não cumprir.

HOLT, “Morton Horwitz and the Transformation of American Legal History”, *William and Mary LR*, 23, 1982, pp. 663-723, e PETER LINZER (org.), *A Contracts Anthology*, s.l., Anderson Publishing, 1989, pp. 129 e ss., com mais indicações.

12 Para não sobrecarregar o texto, omitimos as indicações bibliográficas relativas aos pontos que seguem. Para estas, v. o nosso *Interesse contratual negativo e interesse contratual positivo*, Coimbra, a publicar em 2008, n.º 17 e 18.

A base da análise económica da decisão sobre o cumprimento é a chamada doutrina do “não cumprimento eficiente” (*efficient breach*). Segundo esta, se uma parte não cumpre, mas ainda ficar em melhor situação depois de pagar uma indemnização que compense integralmente a contraparte (de forma a que deixe esta numa situação de indiferença entre receber o cumprimento ou a indemnização), o resultado seria superior em termos de eficiência: consideradas como uma unidade, as partes estarão em melhor situação por causa do não cumprimento e este não deixará ninguém em pior situação (ou, de outra forma, comparando a situação original em termos paretianos, ninguém estará, segundo a sua própria avaliação, em pior situação, e uma pessoa estará, segundo a sua própria avaliação, em melhor situação). O “não cumprimento eficiente” potenciará, pois, um *acréscimo agregado*, de bem-estar social e/ou das partes, levando a evitar o cumprimento indevidamente custoso e permitindo um aproveitamento alternativo (como a venda a terceiro que avaliasse a coisa num montante mais elevado ou uma utilização mais valiosa pelo próprio devedor). O não cumprimento do contrato, com a compensação do credor, poderia ser desejável, por eficiente, nos casos em que os custos para o devedor passaram (por intervenção de um terceiro ou outra alteração dos valores implicados no negócio para o devedor) a superar o ganho para o credor, permitindo ao devedor compensar inteiramente este e ainda ficar em melhor posição do que se tivesse cumprido¹³. A conclusão seria, pois, que, para aumentar a eficiência, a parte que beneficia com o não cumprimento deveria nessa situação *deixar de cumprir*: o devedor deveria ser encorajado a cumprir quando o não cumprimento não fosse eficiente, e encorajado a não cumprir quando o não cumprimento fosse eficiente.

Ora, os proponentes da doutrina da *efficient breach* notaram logo que o não cumprimento tenderá a ser eficiente quando a indemnização for pelo interesse contratual *positivo*, pois nesse caso a parte que pensa em não

13 Noutra perspetiva, podemos perguntar-nos qual seria o contrato que as partes teriam celebrado se tivessem previsto todas as contingências futuras – isto é, um “contrato completo”, que identificasse todas as possíveis hipóteses e especificasse a conduta exigida das partes em cada uma. Ora, nesse caso teriam preferido um contrato completo que liberasse o devedor do cumprimento em certas circunstâncias – pois elas poderiam num tal contrato dividir entre elas o excedente de “bem-estar” criado pelo contrato, e haverá algumas repartições desse excedente em que o ganho para cada parte excede o ganho esperado de um contrato que exigisse o cumprimento em todas as circunstâncias. Ao escolher uma tal repartição cada parte estará em melhor situação *ex ante*, no momento da contratação e disponível para cumprir ou não em todas as situações especificadas, conforme previsto. Um não cumprimento em circunstâncias em que um contrato completo (eficiente) o dispensasse é justamente designado como “não cumprimento eficiente”. Note-se desde já que esta análise requer alguns outros pressupostos, como o de que as partes têm disposições neutras em relação ao risco, que o objetivo de cada uma é maximizar a sua riqueza, que a negociação pós-contratual é impossível (com custos proibitivos resultantes da situação de monopólio bilateral), que o cumprimento parcial não é possível (é “tudo ou nada”) e que os contratos não criam efeitos patrimoniais, ganhos ou perdas, não compensados para terceiros (externalidades).

cumprir apenas o fará se, e só se, o seu ganho exceder o valor do prejuízo *para a outra parte*. O interesse no cumprimento daria aqui precisamente o incentivo necessário para que a ideia do “não cumprimento eficiente” funcionasse: tal interesse (ao obrigar o devedor a “internalizar” todo o prejuízo que o não cumprimento causou ao credor) põe, por definição, a outra parte na mesma posição em que o cumprimento a teria deixado, numa situação em que seja indiferente para ela o cumprimento ou o não cumprimento. Uma indemnização mais elevada, pelo contrário, preveniria o não cumprimento em casos em que seria eficiente, e uma indemnização abaixo do interesse no cumprimento deixaria de prevenir a falta de cumprimento não eficiente¹⁴.

14 A doutrina do “não cumprimento eficiente” ficará melhor esclarecida com a seguinte exemplificação. Em primeiro lugar, um exemplo de “não cumprimento eficiente”. *A* contrata com *B* produzir cadeiras por um preço de 1 000 000, sendo o seu lucro de 200 000. Posteriormente, *C* propõe a *A* que este produza para ele, *C*, mesas por um preço de 2 000 000, sendo o lucro de *A* neste contrato de 750 000, e sendo impossível a *A* cumprir os dois negócios por causa dos prazos de entrega e dos seus limites de capacidade produtiva. Se *A* não cumprir o contrato com *B*, este terá prejuízos de 400 000, compostos por mais 200 000 que *B* terá de pagar a um outro produtor, 100 000 em custos relacionados com o atraso (incluindo quebras de vendas) e 100 000 de custos com a preparação do segundo contrato. Ora *A* pode – e prefere – pagar a *B* os 400 000 que são o prejuízo deste em consequência do não cumprimento (o interesse contratual positivo), e ainda ficar com 350 000 de lucro resultante do contrato relativo às mesas (ou, caso *A* não possa deixar livremente de cumprir, pode negociar com *B* uma indemnização entre 400 000 e 550 000 que deixe ambos em melhor situação do que se o contrato das cadeiras tivesse sido cumprido). O não cumprimento é, pois, eficiente, deixando pelo menos uma parte em melhor situação do que se o contrato tivesse sido cumprido e não deixando a outra em pior situação. Se *A* for obrigado a cumprir o contrato com *B* – mediante uma ação de cumprimento ou outro tipo de sanções mais do que compensatórias –, deixaria de fabricar as mesas para *C*, e seria forçado a não aproveitar a máxima utilidade que os seus recursos (a sua capacidade produtiva) podem produzir, sem benefícios económicos. É esta a razão económica pela qual o não cumprimento não é considerado, no *common law*, como um crime, não são impostas penalidades para o não cumprimento nem se prevê, como regra, a execução específica.

Em contraposição, uma hipótese de não cumprimento não eficiente, induzido pela limitação da indemnização ao interesse negativo. Suponhamos que *A* contrata novamente com *B* a produção de cadeiras por um preço de 1 000 000, e, em vez de qualquer lucro, descobre logo depois da conclusão do contrato que perderá 50 000 com o cumprimento. *B*, que teria de pagar mais 200 000 (isto é, 1 200 000) para obter a produção das cadeiras por outro produtor, não realizou ainda qualquer “investimento de confiança” – isto é, não realizou qualquer despesa nem deixou passar qualquer oportunidade de negócio (pois o melhor contrato alternativo no momento da conclusão do contrato com *A* seria já aquele que exigiria um pagamento de mais 200 000). O interesse na confiança de *B* é, pois, nulo, e, se apenas esse fosse indemnizável, *A* teria um incentivo para não cumprir, evitando um prejuízo de 50 000 e não tendo de pagar nada a *B*. Isto não é, porém, eficiente, pois se *A* cumprir existirá um lucro líquido, isto é, obter-se-á 1 200 000 de valor com um custo de 1 050 000 (note-se, desde já, que a indemnização pelo interesse contratual positivo é que desencoraja esse não cumprimento ineficiente: ao fazer com que *A* tenha de pagar a *B* 200 000, correspondente ao custo de um contrato de substituição, *A* preferirá perder os 50 000 a pagar 200 000, só não sendo assim quando o seu prejuízo ultrapassasse este último valor). Se o custo de um “contrato de cobertura”, em vez de 1 200 000, fosse 1 040 000, o não cumprimento já seria, porém, eficiente, gerando uma afetação mais eficiente da capacidade produtiva de *A*. Observe-se ainda que a eficiência económica seria igualmente servida por uma regra de execução específica (ação de cumprimento), desde que as partes tenham a possibilidade de renegociar. No primeiro exemplo referido, *A* entrará então em contacto com *B* oferecendo-lhe um valor entre 400 000 (o prejuízo de *B* com o não cumprimento) e 550 000 (a diferença entre o lucro para *A* do contrato das mesas e o lucro do

Assim, o interesse contratual positivo incentivaria o cumprimento e o não cumprimento eficientes.

A doutrina do “*não cumprimento eficiente*” não censura, pois, e antes encoraja, o não cumprimento deliberado, voluntário, desde que eficiente, podendo tal eficiência resultar, designadamente, do aparecimento de novas ofertas (o exemplo mais utilizado) ou da alteração da situação, tornando mais onerosa (sem permitir a resolução ou modificação do contrato por alteração das circunstâncias¹⁵) a prestação do devedor, ou menos valiosa, para este, a contraprestação da contraparte.

Só não seria assim nos casos ditos de um não cumprimento “*oportunista*”, isto é, aquele que visa apenas tirar partido de um contexto em que o cumprimento é sequencial, em vez de simultâneo, obtendo os benefícios mas tentando evitar os correspondentes sacrifícios: tal conduta não tem justificação económica e deveria ser simplesmente prevenida, por exemplo, obrigando o inadimplente a restituir todos os benefícios.

Para além do seu fundamento económico (que é também discutido, como referiremos), a emergência e defesa da doutrina do “*não cumprimento eficiente*” foi propiciada por um terreno juridicamente favorável no direito anglo-americano. Segundo uma certa posição, a ordem jurídica não deveria de todo compelir a execução do contrato, e antes deixaria as partes contratantes decidir se querem cumprir o contrato ou, antes, compensar a outra parte pelo não cumprimento. Não é, além disso, necessária no *common law* uma declaração (judicial ou não) de resolução do contrato para o devedor ficar liberado das suas obrigações. Salvo nos casos em que é admissível a execução específica, o não cumprimento da obrigação *in natura* exonera o devedor, ficando este apenas obrigado, em consequência, ao pagamento de uma indemnização (normalmente pela diferença em relação ao preço de mercado ou da operação de substituição), a qual é, pois, a sanção principal, e ficando ainda o credor sujeito a um “dever” (ou ónus) de mitigação dos danos (*mitigation*). E vários outros fatores no direito anglo-americano contribuem também para facilitar o caminho da doutrina do “*não cumprimento eficiente*”¹⁶.

contrato das cadeiras). O ganho de 350 000 resultante do segundo contrato ainda se verificará, portanto, mas a distribuição desse lucro entre devedor e credor será diversa.

15 A teoria do “*não cumprimento eficiente*” não se refere à possibilidade de resolução ou modificação do contrato, por alteração anormal de circunstâncias que constituíram a “base do negócio” em termos de a exigência das obrigações assumidas afetar gravemente a boa fé, prevista entre nós no art. 437.º do Código Civil. Se tal resolução for possível, não se chega a pôr qualquer problema de não cumprimento e de responsabilidade contratual.

16 Nomeadamente: o não cumprimento não tem, para gerar responsabilidade, de ser causado por culpa do devedor; prevalecendo no *common law* na matéria do não cumprimento a abordagem factual sobre uma atitude moralizante, aceitando-se, por um lado, que qualquer negócio importa riscos e quem entra num mercado e opera nele tem de suportar, como natural, o risco do não cumprimento

Ainda que não seja consagrada como regra, a teoria do “*efficient breach*” é hoje amplamente conhecida, mesmo dentro da comunidade jurídica nos Estados Unidos.

Note-se, porém, que a determinação da medida da indemnização pelos tribunais é, nesta ótica económica, uma forma específica do problema de “*revelação de preferências*” das partes, dada a clivagem entre os valores subjetivos e os preços objetivos, e já se tem sustentado que tal problema só parece poder ser resolvido satisfatoriamente atendendo à palavra das partes, tal como resultou da estipulação de cláusulas de liquidação antecipada da indemnização. Ora, a favor da *efficient breach*, pergunta-se por que razão as partes não escolhem sistematicamente medidas de indemnização tão elevadas que lhes garantam o cumprimento, respondendo-se pondo em causa que realmente prefiram indemnizações supracompensatórias, já que tais indemnizações induziriam em certos casos o cumprimento do contrato mesmo que não fosse eficiente. Uma regra de estrito *pacta sunt servanda*, sem exceções, induz incentivos para cumprir *sempre* contratos, que podem porém conduzir a uma afetação ineficiente de recursos, o que não interessará às partes.

A doutrina do “não cumprimento eficiente” pode, por outro lado, ser vista como uma forma de prolongar, mesmo depois do momento da conclusão do contrato e até à decisão sobre a realização fáctica do seu cumprimento, a *afetação de recursos pelo mercado*, para os seus usos mais valiosos, com a continuação da operatividade do mecanismo da concorrência entre as utilizações alternativas dos bens ou serviços objeto do contrato – incluindo entre o credor e um terceiro cuja proposta apareça posteriormente. Se a afetação de recursos puder ser melhorada com o não cumprimento do contrato e sem desvantagem para ninguém, isto é, se existe uma forma de conseguir uma melhoria (designadamente, em termos paretianos), tal doutrina recomenda o não cumprimento.

É claro que tal entendimento do não cumprimento – e das medidas de indemnização como incentivos a ele – contraria, por outro lado, a ideia de que ele seria um ato *imoral*, por significar a *quebra de uma promessa*, contestando-se a caracterização típica e a avaliação em termos morais da situação de “não cumprimento eficiente”, e notando-se que um “*contrato completo*” preveria, para tal

pelas suas contrapartes, e, por outro lado, que quem muda de intenções e vende os bens a uma melhor oferta, apesar do seu não cumprimento, simultaneamente cumpre as obrigações para com o terceiro (embora o alargamento da *tortious interference with contractual relations* tenha vindo a pôr limites a estas possibilidades como mencionaremos); à parte que não está disposta a correr riscos resta exigir, antes de contratar, a inclusão de cláusulas de liquidação antecipada da indemnização, ou, até, com certos limites, de *penalty clauses*.

situação, um não cumprimento, resultando a manutenção da vinculação, nesse caso, apenas da “incompletude”¹⁷.

Em teoria, a *efficient breach* promoveria, pois, um resultado eficiente nos contratos de execução diferida – diversamente de uma orientação que admita, por exemplo, a execução específica mediante uma ação de cumprimento, a restituição dos ganhos obtidos pelo inadimplente, a responsabilidade do terceiro que cause o inadimplemento, os *punitive damages*, ou, em geral, quaisquer sanções mais do que compensatórias para o credor.

3. Crítica à teoria do “não cumprimento eficiente”

A um olhar mais atento, porém, logo se conclui que a teoria da “*efficient breach*” deve ser decididamente *rejeitada*. Ela tem, aliás, sido submetida a forte crítica, subsistindo na doutrina – mesmo entre aquela que olha para o Direito de uma perspectiva económica – o desacordo sobre a sua validade. Tal doutrina, ainda que fosse correta do ponto de vista da ciência económica, não deve ser utilizada para extrair conclusões *normativas* no plano jurídico sobre o melhor regime a adotar – designadamente, para tentar concluir que a consequência jurídica ótima é, nas hipóteses a que se refere, a exoneração do devedor da obrigação de cumprimento.

Antes de mais, observa-se que, também aqui, o *estatuto* de tal teoria não é muito claro, oscilando entre a *descrição* no plano económico e a *prescrição* normativa, pretensamente jurídica. Parece que podem distinguir-se duas pretensões em tal teoria: por um lado, uma pretensão descritiva do direito vigente sobre a indemnização contratual (variando, evidentemente, consoante a ordem jurídica); por outro lado, uma pretensão (normativa) de ser o fundamento para decisões de casos duvidosos ou uma regra geral a adotar *de jure condendo*.

No entanto, nem uma nem outra pretensão dessa doutrina podem considerar-se procedentes, quer entre nós quer no *common law*, onde os tribunais a não têm acolhido como regra¹⁸, levantando a teoria muitas dúvidas também no próprio plano económico.

Quanto à incompatibilidade com o *regime jurídico* do não cumprimento das obrigações, ela resulta, claramente, da regra do *cumprimento pontual* dos

17 Num sentido económico, e não jurídico, pois não se pretende que exista qualquer verdadeira lacuna contratual.

18 V., aliás, CRAIG S. WARKOL, “Resolving the Paradox Between Legal Theory and Legal Fact: The Judicial Rejection of the Theory of Efficient Breach”, *Cardozo LR*, 20 (1998), pp. 321-353, perguntando-se por que razão os tribunais falharam na aplicação da teoria da *efficient breach*, que considera tão solidamente firmada na ciência económica.

contratos (artigo 406.º do Código Civil Português¹⁹) e do facto de a ação de cumprimento e a *execução específica* não serem vistas, na grande maioria dos sistemas continentais, como “remédios” residuais, mas antes como a regra (cf. os artigos 817.º e seguintes), afirmando-se uma “prioridade natural e jurídica” do cumprimento e da indemnização em forma específica (cf. o artigo 566.º, n.º 1²⁰).

É certo que a possibilidade de cumprimento conhece várias limitações, não só jurídicas como fácticas – económicas, materiais e até sociais ou psicológicas –, e que a indemnização (por equivalente) será, provavelmente, em termos estatísticos, o “remédio” mais frequente para o não cumprimento da obrigação contratual. Todavia, nos sistemas jurídicos continentais a *regra jurídica* é, sem dúvida, a da exigibilidade do cumprimento *in natura*, sem qualquer limitação do credor a uma indemnização, ou qualquer concessão de uma faculdade alternativa ao devedor, de optar por esta última. Com base na supremacia “natural e jurídica” do cumprimento, há até na nossa doutrina quem defenda a proibição da renúncia antecipada, unilateral ou por convenção²¹, ao direito de recorrer à execução específica para obter a própria prestação devida (embora não nos pareça essa a posição preferível).

Compreende-se, pois, que as ordens jurídicas continentais optem por censurar de forma mais estrita o não cumprimento *voluntário* do que o *meramente culposo*, e sem tomar em conta (pelo menos expressamente) considerações de eficiência do não cumprimento. É o que acontece, em várias ordens jurídicas, com normas que alargam o âmbito dos direitos do lesado – em particular dos danos ressarcíveis – quando o não cumprimento for intencional²².

Há várias razões para essas opções pela prioridade do cumprimento, e pela tentativa de tornar o não cumprimento um evento excepcional, a ser evitado e que não é justificado pela alteração das expectativas da parte que não cumpre.

19 Diploma a que pertencem todas as normas citadas sem indicação especial.

20 Pomos aqui entre parêntesis o problema de saber se esta norma se aplica à responsabilidade por não cumprimento do contrato (nomeadamente, quando este resulta da conversão da mora em não cumprimento), tendo o credor optado já pela indemnização (cf. o art. 808.º, n.º 1), ou se deve ser excluída essa aplicação, para não permitir um “renascimento” da pretensão de cumprimento pela via da reconstituição natural. A resposta a esta questão parece depender, além do mais, de saber qual é o efeito do art. 808.º, n.º 1, sobre a pretensão de cumprimento

21 A cláusula pela qual, se o devedor preferir não cumprir, o credor apenas poderá exigir uma indemnização por não cumprimento, e não a execução específica, distingue-se, por outro lado, da chamada multa penitencial, que permite ao devedor liberar-se da obrigação pagando a “multa” previamente fixada, desde logo, justamente porque na primeira o montante (liberatório) está previamente fixado por convenção das partes.

22 V., entre nós, por ex., o artigo 891.º sobre a resolução do contrato, e, embora para o cumprimento defeituoso (e discutindo-se o sentido do termo “dolo” nalgumas dessas normas), os arts. 898.º, 908.º, 957.º e 1134.º Os efeitos de tal censura em geral são bem mais amplos nos direitos francês e italiano, em que os arts. 1150.º e 1225.º, resp. do *Code Civil* e do *Codice Civile*, preveem que o limite da previsibilidade dos danos não se aplica se o devedor tiver procedido com dolo.

Refira-se, assim, a necessidade de sublinhar a *força vinculativa do contrato*, a fundar a *certeza do cumprimento* nas relações contratuais (caso contrário, estar-se-ia sempre sujeito a um possível não cumprimento). No mesmo sentido aponta, também, a necessidade de assegurar o processamento com correção do *tráfego comercial*, já que as partes avessas ao risco de potenciais não cumprimentos poderiam desde logo evitar a conclusão de contratos sem o necessário grau de certeza. Invoca-se, ainda, a necessidade de *moralizar* a atuação no mercado, tendo os ditames da boa fé como padrão a que os operadores devem conformar-se (padrões moralizadores, aliás, cuja interiorização levará, por sua vez, a um efeito desonerador da formação de expectativas individuais e da criação de incentivos concretos ao cumprimento, pois o afastamento deles será visto como um custo, subjetivamente e em termos de reputação). E, diz-se, também, que a redução de certos *custos da transação* é conseguida pelo reforço do direito ao cumprimento, já que as partes deixam de, na perspetiva do potencial não cumprimento (unilateral ou mútuo), ter de conseguir obter garantias ou segurança, seja por mecanismos contratuais de planeamento e negociação mais custosos (como cláusulas penais, cláusulas resolutivas, etc.) seja mediante a contratação de garantias bancárias e apólices de seguro (também com custos evidentes).

O nosso direito opta, pois, claramente pela prioridade do cumprimento, sem que apresente quaisquer traços da relevância da eficiência do cumprimento – não sendo esse o caso, designadamente, das normas sobre a fixação da indemnização em dinheiro. A comprová-lo, estão também, numa das hipóteses mais frequentemente apontadas de “não cumprimento eficiente” (a da oferta superveniente de terceiro), as propostas de *responsabilização do terceiro* que contrata com o devedor pelo não cumprimento da obrigação (pois parecem não ter ainda obtido um acolhimento generalizado ou expresso reconhecimento *de jure condito*), de eventual extensão da responsabilidade civil aos *ganhos obtidos* pelo violador do contrato (consagração de uma *Gewinnhaftung*, ou “responsabilidade pelos ganhos”, incompatível com a configuração da responsabilidade civil à volta de um dano, que, no caso, o credor não sofre), ou do reconhecimento de uma obrigação de restituição ou *remoção (“disgorgement”)* dos lucros ou benefícios obtidos pelo credor com o inadimplemento²³. Esta obrigação, se não parece ter consagração específica num dos tipos normalmente admitidos de enriquecimento sem causa, existirá já hoje, sem dúvida, pelo menos, quando

23 O problema tem sido discutido sobretudo nas ordens jurídicas que não reconhecem em geral o direito à execução específica do contrato, como acontece no *common law*.

tais benefícios se traduzirem na aquisição de um direito sobre certa coisa, ou contra terceiro, mediante o instituto do *commodum* de representação (artigos 803.º e 794.º²⁴).

As ordens jurídicas continentais não revelam, assim, *qualquer preferência* pelo não cumprimento, mesmo quando os benefícios para o credor são superiores ao custo – ainda que de oportunidade – para o devedor.

Aliás, mesmo no direito anglo-americano parece haver poucas razões para se aceitar que a doutrina da *efficient breach* realmente é a base de quaisquer regras vigentes sobre a indemnização contratual.

Tal doutrina concilia-se mal com muitos aspetos do direito vigente, não sendo correta uma afirmação geral no sentido de que esta não quer evitar o não cumprimento eficiente. Os tribunais tomam em conta a *voluntariedade* da violação do contrato como agravante, designadamente, alargando o padrão da certeza exigida para a prova dos lucros do lesado (para além de ser considerada para outros efeitos, como, por exemplo, a importância do não cumprimento, como *material breach*). Além disso, o *tort of interference with a contract* é evidentemente inconciliável com a ideia de que o sistema jurídico aprova a violação eficiente do contrato, pelo menos, nos casos em que ela resulta da intervenção de terceiros com melhores propostas. Em muitos casos nota-se também, atualmente, uma tendência no *common law* para alargamento do âmbito da *execução específica*, com os tribunais crescentemente dispostos a decretá-la e os legisladores cada vez mais dispostos a prever tal possibilidade, tendência, essa, que também conflua com a doutrina do não cumprimento eficiente. Em alternativa, existem também vozes no sentido de que as violações intencionais do contrato sejam prevenidas mediante o recurso aos *punitive damages*, reforçando o direito do credor. E, por último, os tribunais do *common law* têm nalguns casos – embora não como regra geral – exigido a remoção ou *disgorgement* dos lucros que o inadimplente obteve, como consequência do não cumprimento, mesmo que as partes não estivessem ligadas por qualquer relação fiduciária, existindo um amplo debate sobre a aceitabilidade de tal extensão das hipóteses de obrigação de restituição (extensão dos, aliás impropriamente, chamados “*restitutionary damages*”)²⁵.

24 Apesar de o art. 803.º, n.º 2, mandar reduzir o montante da indemnização na medida correspondente ao valor do direito adquirido pelo devedor contra terceiro, não existe qualquer norma que expressamente limite a obrigação de entrega do *commodum* de representação (incluindo o *commodum ex negotiatione*) ao valor dos prejuízos causados ao credor pelo não cumprimento. A questão da limitação do direito ao *commodum* de representação ao montante da indemnização do credor pelo interesse positivo (pelo valor da prestação) foi, porém, discutida na doutrina alemã.

25 A discussão foi renovada com o caso *Attorney Gen. v. Blake*, 4 All E.R., p. 385, de 2000 (*House of Lords*), em que foi ordenada a restituição dos lucros que um antigo espião obtivera com a publicação de um livro em violação do contrato que havia celebrado com a Coroa britânica. Mas o caso suscita dúvidas como precedente, pois, embora não para a questão de saber se o demandante tinha um direito aos

Também a pretensão normativa da doutrina do “não cumprimento eficiente” é de considerar improcedente, quer como regra integradora de lacunas quer como proposta geral *de jure condendo*.

Pode duvidar-se da verificação, no mundo real, de alguns dos *pressupostos* simplificados em que tal doutrina assenta, e que a levam a ignorar não só os limites ao apuramento de valores subjetivos (e de danos não patrimoniais), como a ter mesmo de recorrer a comparações com base no critério da disponibilidade para pagar.

Entre tais limitações reais dos pressupostos da teoria encontram-se, desde logo, os desvios em relação à pura racionalidade económica, que se verificam sistematicamente nos atores no mundo real e que se encontram bem comprovados e procuram ser considerados pelas correntes “behavioristas” da análise económica – desde a “racionalidade limitada” resultante de limites temporais aos limites de cognição e previsão (e o recurso a mecanismos heurísticos e “regras de decisão” gerais), até ao conhecido “efeito de dotação”²⁶. Perante tais *distorções* – e os problemas de informação de muitos agentes sobre as consequências que as suas condutas acarretarão – poderá em muitos casos duvidar-se da amplitude e linearidade do efeito indutor de comportamentos das regras que preveem consequências indemnizatórias pelo não cumprimento.

Em todo o caso, enquanto não se encontre um sistema que prove ser superior na explicação das consequências das normas jurídicas e de outras influências sobre o comportamento humano (incluindo as emoções²⁷), pode aceitar-se que a hipótese da racionalidade individual, do ser humano racional, egoísta e

lucros obtidos pelo demandado, invocou-se o seu “interesse legítimo em prevenir esta atividade lucrativa do demandado, e, portanto, em privá-lo do seu lucro”, interesse, esse, que se ligava à necessidade de preservar a confiança dos informantes e de manter a moral dos funcionários dos serviços secretos, mesmo quanto a informação (como a que estava em causa) que já não era confidencial.

26 O *endowment effect* é o efeito que se exprime no facto (experimentalmente comprovado) de o montante que as pessoas exigem para abrir mão de um bem ser maior do que aquele que estão dispostas a oferecer para o adquirir – ou, por outras palavras, que resulta de se avaliar em mais o que se tem do que o que se não tem, e, portanto, de a exigência para receber por um bem ser (bastante) maior do que a disponibilidade para pagar por ele. Uma das explicações do “efeito de dotação” reside na “preferência pelo *status quo*”: as pessoas teriam uma função de utilidade em forma de “S”, primeiro convexa e depois côncava, com inflexão justamente na “dotação” (*status quo* ou “aversão às perdas”), o que explicaria que exigissem bastante mais para abrir mão de algo do que pagariam para obter a mesma coisa. A preferência pelo *status quo* resultaria de as desvantagens de perder algo se perfilarem muito maiores do que as vantagens de o adquirir.

É interessante notar que algumas regras jurídicas traduzem justamente esta preferência pelo *status quo* (mesmo que a “linha de base”, que separa, seja por vezes difícil de determinar com precisão), como quando se prefere o efeito *de damno vitando* ao *de lucro captando*.

27 V. ANTÓNIO R. DAMÁSIO, *Descartes’ Error. Emotion, Reason and the Human Brain*, New York, Grosset/ Putnam, 1994 (trad. port. *O Erro de Descartes. Emoção, Razão e Cérebro Humano*, Mem-Martins, Publicações Europa-América, 2000).

maximizador de utilidade – a hipótese *REMM* (*resourceful evaluating maximizing man*) ou REM (de *rationales egoistisches Mensch*), não tem de ser inteiramente posta de parte, como mero modelo explicativo.

A estas limitações acresce, porém, uma crítica “*subjetivista*”. O “não cumprimento eficiente” pressupõe a possibilidade de medir o valor do interesse do credor no cumprimento, como equivalente monetário que deixaria o credor indiferente entre ele e o cumprimento, sendo que, evidentemente, para tal “indiferença” não é bastante apenas o valor normal, objetivo, do cumprimento. Proteger o interesse *subjetivo* da contraparte no cumprimento é um elemento central na noção de “não cumprimento eficiente”, tendo o inadimplente de reparar as perdas reais do lesado, incluindo os custos que teriam sido evitados sem o cumprimento. Segundo a lógica dessa doutrina, o contratante que pense em não cumprir deve tomar em conta mais do que o valor objetivo do cumprimento para a contraparte e proceder ao cálculo do “lucro líquido”, esperado a partir de uma recusa de cumprimento do contrato, com base numa projeção de todos os custos – isto é, dos custos normais do mercado mais as despesas do negócio e as causadas pela necessidade de resolver o litígio pelos meios normais – necessários para pôr o lesado, subjetivamente, numa posição de indiferença entre o cumprimento e a indemnização. Este ponto é crucial, sendo que, se o interesse no cumprimento for demasiado baixo, existirá demasiado não cumprimento, e se for excessivamente elevado, segundo esta doutrina induzir-se-ia em excesso ao cumprimento, em ambos os casos com resultados não eficientes.

Ora, pode acontecer – e acontece frequentemente – que o credor atribua um valor ao cumprimento *diverso do valor de mercado*, devido à sua posição subjetiva ou à sua estrutura de preferências (também económicas, e não necessariamente apenas resultantes de um valor afetivo). Sob pena de violação dos postulados em que assenta a própria análise económica, a autonomia e as preferências dos sujeitos em causa não podem então ser sacrificados a uma avaliação objetiva, pelo mercado ou por uma autoridade, do interesse no cumprimento, a qual poria em risco mesmo a eficiência (pois o sujeito cujas preferências não forem adequadamente refletidas ficaria prejudicado em termos de bem-estar).

Isto torna, assim, incontornável o problema do método para o devedor conseguir determinar o valor subjetivo antes de decidir – o problema conhecido da “*revelação de preferências*” do devedor e do credor, cuja solução pode ter custos muito elevados ou ser mesmo impossível (designadamente, na falta de estipulações das quais se deduzam, como cláusulas de liquidação antecipada da indemnização). Muitas vezes estas *limitações e custos de apuramento da informação* (os custos de informação) não são sublinhados na exposição da

doutrina do “não cumprimento eficiente”, mas é claro que eles podem, só por si, inviabilizar a aplicação de tal doutrina ou (na própria lógica económica) torná-la ineficiente.

Para tornar esta dificuldade, como critério para a revelação do valor subjetivo, a análise económica recorre à *disponibilidade para pagar*, partindo do princípio de que o valor subjetivo será tanto maior quanto maior for o preço oferecido por uma pessoa por um bem. Parece claro, porém, que a aplicação deste critério (embora torneando o problema da insinceridade nas preferências declaradas) não é (pelo menos sempre) suscetível de revelar o valor subjetivo: apenas pode servir para medir o valor atribuído conjugado com a capacidade para pagar certo preço, mas não a possibilidade de pagar pelo facto de se não dispor (nem ter possibilidade de acesso) ao montante que traduza o valor atribuído ao bem (de se não *poder pagar*...).

No próprio plano da avaliação dos custos e benefícios, surgem, pois, logo incertezas práticas sobre a eficácia relativa da violação do contrato ou do seu cumprimento no mundo real. E se, por exemplo, o terceiro que oferece um preço tiver avaliado em termos excessivamente favoráveis o bem, a doutrina do “não cumprimento eficiente” conduzirá logo a uma afetação menos eficiente de recursos.

Acrescem as dificuldades que a incorporação de uma *terceira alternativa* (em relação ao dilema entre cumprimento e não cumprimento) necessariamente acarreta: a possibilidade de *renegociação* do contrato. Na verdade, um modelo que pressuponha que as partes não são capazes de comunicar entre elas e negociar antes do (não) cumprimento é necessariamente reductor. A análise económica tentou incorporar a possibilidade da renegociação, com resultados aparentemente contraintuitivos.

Aplicando ao problema da *renegociação* o “teorema de Coase” concluiríamos que, sem custos de transação, as partes seriam levadas, mesmo sem qualquer doutrina de “não cumprimento eficiente”, a negociar e afetar entre elas o direito ao cumprimento (ou a não cumprir) àquela que lhe atribua mais valor (tendo compensado a outra parte). A preferência pelo “não cumprimento eficiente” ignora, porém, num tal contexto, que, num mundo coaseano, mesmo depois do não cumprimento, não existe diferença, em termos “alocativos”, entre a previsão legal do cumprimento sem limitações e a do não cumprimento eficiente. As teorias sobre este seriam, pois, *irrelevantes*, tal como a regra jurídica a esse respeito: as partes simplesmente negociariam de tal modo que o direito ao (não) cumprimento acabe na titularidade de quem o avalia mais (em termos de possibilidade de pagar, repete-se). Mas isto significa também que não se poderia utilizar a análise económica, do funcionamento do mercado, para concluir qual é a teoria mais correta, pois o resultado daquele é o mesmo *seja qual for a*

teoria. Se, num tal mundo, a riqueza não aumentaria nem diminuiria por força da regra jurídica do cumprimento, a única (mas juridicamente fundamental) diferença, entre o direito ao não cumprimento e a exigência de uma renegociação (pela manutenção do direito ao cumprimento), estaria em que, no caso de renegociação (ou na ação de cumprimento), o credor iria participar no benefício relativo do devedor, sendo partilhado entre ambos, ou, mesmo, apropriado apenas pelo credor. Ou seja, o que é *economicamente irrelevante* é o que é *juridicamente fundamental*: a atribuição do benefício resultante do (não) cumprimento (como, aliás, acontece também noutros exemplos de aplicação da análise coaseana, relevante em termos de eficiência alocativa, mas sem nada a dizer para a alteração de resultados no fundamental problema “distributivo” – *rectius*, de justiça –, cuja avaliação é constitutiva de uma abordagem de tipo jurídico).

Tal reflexão deixa, também, ver que, num *mundo com custos de transação*, a eficiência do não cumprimento não pode ser avaliada sem um conhecimento pormenorizado desses custos de transação e sem a inclusão de custos como os da renegociação ou do litígio entre as partes, e ainda (não só no caso de litígio) dos já referidos custos de informação do devedor sobre os benefícios produzidos pelo cumprimento do contrato, em comparação com a alternativa que se lhe oferece.

Voltando ao aspeto juridicamente mais importante, nota-se que o direito ao cumprimento e a doutrina do “não cumprimento eficiente” conduzem a *resultados “distributivos”* entre as partes – isto é, de justiça contratual – bastante *diferentes*. Se o não cumprimento não for permitido, o novo excedente é, no mínimo, partilhado pelo credor (ou entre este e um terceiro que faça melhor oferta), enquanto num regime de “não cumprimento eficiente” tal excedente será apropriado apenas pelo devedor, e, quando muito, dividido entre este e o terceiro. Se os custos de informação não forem proibitivos, não se vê, porém, por que razão a reafetação de recursos tem de ocorrer através do não cumprimento, e não pode verificar-se antes *por intermédio do credor*, pelo menos nos casos (que serão a maioria) em que a alteração do excedente produzido com o contrato em comparação com afetações alternativas não impede que seja o credor a proceder à reafetação – por exemplo, sendo o credor, e não o devedor, a vender ao terceiro²⁸. Ou seja, não se vê por que razão o ganho superveniente (da entrada de um terceiro, ou da alteração das condições de mercado) não é, como resultaria do acordado pelas partes, apropriado *pelo*

28 Assim, tb. se o aumento do valor do cumprimento, e da onerosidade para o devedor, resulta de uma alteração no mercado. Em certos casos, pode acontecer, porém, que os custos acrescidos para o devedor não se traduzam num aumento do valor da prestação (ou que o valor da contraprestação diminua apenas para devedor, que a receberia), e que apenas o devedor tenha agora conveniência na reafetação do recurso a uma utilização diversa do cumprimento.

credor, que poderia negociar com o terceiro ou vender o bem no mercado, ou, simplesmente, aumentar o valor do seu património, mas antes pelo devedor (isto, sendo certo que, normalmente, as condições que legitimariam um não cumprimento, como o maior valor do bem ou a entrada em cena de um terceiro, também fariam aumentar o valor da prestação para o credor²⁹). A justificação, por vezes avançada, segundo a qual tal revenda pelo credor aumentaria os custos porque implicaria uma *transação adicional* afigura-se-nos improcedente, designadamente, em comparação com os custos, implicados numa solução de não cumprimento, da liquidação do primeiro negócio, incluindo os custos do litígio sobre o cumprimento.

Também aqui, aliás, a raiz da doutrina do “não cumprimento eficiente” – ou pelo menos uma sua importante condição de possibilidade – parece residir, ou numa *sobrevalorização do momento da execução fáctica* do contrato (e na desvalorização da eficácia jurídica deste) ou no já apelidado “perene passo em falso” quanto ao alcance e à correspondente *força vinculativa da promessa*. Trata-se, mais precisamente, da ausência de reconhecimento de que o direito ao cumprimento não tem subjacente uma mera “regra de responsabilidade” (uma “*liability rule*”), antes o contrato atribui, em regra, ao credor (pelo menos entre nós) um direito fundado numa “*property rule*”³⁰, de que o titular pode abrir mão (não é um direito inalienável) negociando um resultado igualmente eficiente, mas com uma bem diversa distribuição patrimonial, em termos correspondentes ao acordado. O que equivale, em suma, ao falhanço do reconhecimento de que o contrato visa normalmente a *transmissão para o credor do direito aos ganhos* ou lucros que resultem da execução, ainda que eles venham a ser aumentados

29 Na ausência de uma “alteração das circunstâncias” que legitime a resolução ou modificação do contrato. Não referimos a ausência de lacunas porque juridicamente a função normal do mecanismo contratual é, justamente, a repartição entre as partes do risco de alterações de valores (do mercado, da conjuntura, etc.), e, portanto, deve manter-se vinculante mesmo nessas hipóteses, e ainda que a hipótese de alteração não fosse especificamente prevista. Bem diversa de tais lacunas contratuais é a “incompletude” do contrato em sentido económico: enquanto nas primeiras o plano que exigiria a previsão (o *Soll/seite*, ou lado de “dever-ser”, do conceito de lacuna) é obtido, designadamente, a partir das necessidades práticas para a execução do contrato com consideração das normas supletivas (que em princípio não são afastadas por uma mera vontade hipotética ou conjectural), na última o critério é, antes, a hipótese teórica (mecanismo ou ideia apenas metodológica ou teoricamente operativo) de uma especificação integral, pelas partes, de todas as possíveis contingências que afetem o contrato e sua execução. Trata-se aqui, no “contrato completo” da teoria económica, de um conceito puramente ideal, com finalidade heurística, sem correspondente na realidade social. Assim, um contrato pode bem ser completo do ponto de vista jurídico, sem lacunas, mas estar longe do “contrato completamente especificado” da teoria económica.

30 Referimo-nos à conhecida distinção de GUIDO CALABRESI/DOUGLAS A. MELAMED, “Property Rules, Liability Rules, and Inalienability: One View of the Cathedral”, *HLR*, 85, 1972, pp. 1089-1128. Sobre a distinção de regras, por CALABRESI e MELAMED, v., agora, FRANK I. MICHELMAN, “There Bave to Be Four”, in *Maryland LR*, vol. 64 (2005), pp. 136-158.

por acontecimentos supervenientes, como o aparecimento de ofertas de terceiros.

E diga-se ainda, por fim – mas não por último –, que há sem dúvida *outros valores* que o direito dos contratos serve para além da promoção da “eficiência de alocação”, alguns dos quais até tornam possível o próprio desempenho de funções económicas do contrato.

Desde logo, a ideia de que a doutrina do “não cumprimento eficiente” deveria ser difundida para encorajar os contratantes a não cumprir afetaria, sem dúvida, a confiança e o crédito, e, mesmo, a possibilidade de certos mercados, como os de mercadorias. As empresas carecem de ter a certeza da percepção do preço que negociaram. Elas reduzem mesmo tal incerteza por operações “de cobertura” (*hedging*) nesses mercados. Ora, tal garantia contra o risco de incerteza do preço, pelo *hedging*, seria impossível se o vendedor de mercadorias pudesse violar um contrato de opção livremente, simplesmente por ter encontrado alguém que oferecesse um preço melhor, sendo que, se a jurisdição em causa não contém uma regra de “não cumprimento eficiente”, a reação normal será então a previsão, pelas partes, de uma cláusula de liquidação antecipada da indemnização que permita ao credor manter tal certeza – o que, como é evidente, também afetará a amplitude da aplicação da doutrina do “não cumprimento eficiente”.

Por outro lado, a concentração nos aspetos pecuniários do não cumprimento não toma em conta aspetos *qualitativos*, eventualmente muito relevantes, e mesmo o valor *moral* do cumprimento da palavra dada, para além de outras sanções não económicas ou jurídicas, como as consequências em termos de reputação, do não cumprimento do negócio. Não é aqui necessário determinar o fundamento do sentimento moral de fidelidade à palavra dada (religioso, jusnaturalista, humanista, teórico-discursivo, etc.). O certo é que o contrato, segundo a experiência prática que dele tem a grande maioria das pessoas, e que está subjacente à lei, funda uma obrigação que funciona, justamente, para isolar a decisão de cumprimento de um cálculo global de valores alternativos (ou da interferência de outras razões), de tal forma que a vinculação (e o respetivo sentimento) se mantenha, com exclusão de outras considerações, ainda que o cumprimento não sirva para, globalmente, evitar um mal, maximizar a utilidade social ou promover as melhores consequências em geral.

Tal como a doutrina do “não cumprimento eficiente” apresenta a sua fundamentação, não existiria, aliás, razão para não admitir, também, uma “*apropriação eficiente*” de recursos alheios, mesmo fora de um contrato – um “*furto eficiente*” ou um “*abuso de confiança eficiente*”. E pode mesmo dizer-se que, ao recusar em geral um fundamento moral para a ideia de que *pacta sunt servanda*, a doutrina do “não cumprimento eficiente” põe em causa a *confiança*, o

valor da *certeza* de que o devedor cumprirá, e o próprio *crédito*: uma prosse-
cução da máxima eficiência no curto prazo destrói a eficiência no longo prazo.

4. Os efeitos dos “remédios” indemnizatórios na responsabilidade contratual

A referida rejeição da doutrina do “não cumprimento eficiente” não impede, porém, que possamos suspender por um momento tal posição e ver resumidamente a que resultados pode levar (ou tem levado) a análise económica da indemnização contratual, vista como incentivo aos contraentes.

É claro – note-se desde já – que tal perspectiva abandona, em rigor, o terreno das verdadeiras “*consequências jurídicas*” de um evento lesivo. A intenção é, antes, analisar e projetar os *incentivos* adequados, para obter diretamente os resultados mais eficientes da conduta das partes em diversas “margens”, ou, mediatamente, para diminuir custos relacionados com o não cumprimento (de informação, de litigância, etc.). A ideia de que os “remédios” para o não cumprimento são, na realidade, incentivos para as partes significa, não que visem uma *compensação*, ou reparação, do prejuízo causado no passado – por exemplo, de acordo com um critério de causalidade baseado na ideia de justiça corretiva, e à luz dos fins das normas que preveem a reparação –, mas antes que visam produzir efeitos, seja anteriormente, pela sua antecipação pelas partes, seja no futuro. Não é de admirar, assim, que se proponha uma “reconsideração do princípio da compensação no direito dos contratos”, e que se saliente que a análise económica do direito dos contratos “dificilmente menciona a compensação, seja como um fim seja como um meio para maximizar o bem-estar agregado das partes contratantes”.

Por outro lado, note-se, preliminarmente, que o problema da medida da indemnização se cruza, também, com o do critério para a determinação das *regras supletivas (default rules)*, não acordadas pelas partes, para a indemnização nos chamados “*contratos incompletos*” (contrapostos aos contratos completos ou “totalmente especificados”), por não especificarem de modo esgotante as soluções para todas as situações possíveis – sendo, porém, igualmente propostas outras abordagens ao problema da análise da medida da indemnização³¹.

31 Para uma delas, o problema seria o de conceber medidas de indemnização que induziriam as partes a tomar a mesma decisão sobre o cumprimento, ou sobre a confiança anterior ao cumprimento, que teriam tomado se não fossem sujeitos separados, mas divisões de uma única empresa integrada tendo o único objetivo de maximizar o valor desta. Outra, por sua vez, considera a indemnização contratual como uma opção de compra do direito ao cumprimento, por parte do inadimplente, a qual expira no prazo fixado para o cumprimento e tem como preço de exercício o montante da indemnização, que se reflete (como o valor da opção) no preço previsto no contrato.

Como dissemos, a medida da indemnização devida pelo inadimplente foi primeiro vista pela análise económica na perspectiva da optimização da *decisão de cumprimento* – isto é, da obtenção do grau óptimo de (não) cumprimento (de um cumprimento e não cumprimento eficientes), tendo passado a ser doutrina comum a de que tal incentivo se atingiria com a indemnização do interesse positivo como criador de incentivos ao cumprimento eficiente.

A previsão de uma indemnização pelo dano de não cumprimento seria o incentivo correto para tal, pois obriga o inadimplente a considerar, “internalizando”, os custos que o seu não cumprimento inflige à outra parte, dando-lhe, conseqüentemente, o incentivo para tomar uma decisão de não cumprimento eficiente, além de o incentivar a tomar precauções contra eventualidades que ameacem aumentar o custo do cumprimento. A indemnização pelo interesse positivo tenderia, pois, a aumentar o valor agregado que o negócio oferece a promitente e promissário, o que justificaria a sua previsão, não só por esse aumento em si mesmo, como, para as partes, porque estas, de um ponto de vista *ex ante*, teriam razões para estar de acordo com uma regra que *maximizasse o excedente* contratual total (o “bolo contratual”), disponível para a divisão entre elas mediante negociação. E à mesma conclusão se chega com outras abordagens da medida de indemnização em termos de eficiência, desde que se considere apenas a decisão de cumprir ou não cumprir (ou, também, a decisão sobre o nível de precauções do devedor contra o risco de inadimplemento, ou de aumento do custo do cumprimento)³².

Apesar dos possíveis efeitos benéficos da regra da indemnização pelo interesse positivo, a análise económica permitiu salientar a *insuficiência* de uma tal conclusão, identificando várias considerações que apontam contra a eficiência dos resultados produzidos por uma regra geral do interesse contratual positivo.

Numa breve ilustração de tais desafios à regra de indemnização pelo interesse contratual positivo – e contra a ideia de uma indemnização compensatória pelo não cumprimento –, nota-se, como já referimos, que existem sérios problemas de obtenção e de custo da *informação necessária* para avaliar completamente o interesse no cumprimento, podendo mesmo as partes ficar numa posição melhor com uma indemnização não tão elevada (não compensando inteiramente o lesado) mas que reduza o (valor e o) risco da *litigância*, apontando-se alguns mecanismos legais que permitem justamente uma tal fixação da indemnização “infracompensatória”.

32 E à mesma conclusão se chega ainda se se considerar a função do direito dos contratos de induzir a cooperação em jogos não cooperativos, pela alteração do *pay-off* de cada jogador.

E, sobretudo, nota-se que a limitação ao incentivo ao cumprimento (mesmo sem considerar outros problemas, relativos ao valor no tempo da opção de não cumprir, quando ambas as partes dispõem de tal opção) é excessivamente *simplicista*, pois a decisão de cumprir é apenas *uma das decisões* que as partes tomam e que pode afetar o excedente contratual – além da decisão de confiar, ou fazer despesas de confiança, de tomar precauções, de tentar mitigar os danos, de procurar uma contraparte, de gastar com a preparação e conclusão do contrato, de concluir o contrato, etc. –, e se a indemnização do interesse positivo induz à optimização daquela decisão, os incentivos que dela resultam marginalmente para estas outras decisões parecem em grande medida afastar-se da eficiência.

Assim, a indemnização do interesse positivo afeta a *decisão do credor de confiar*, de realizar “investimentos de confiança” (que podemos aqui ver quer como despesas quer como a renúncia a outras oportunidades), em termos não eficientes: induz *excesso de confiança* no credor, o qual, devido a essa indemnização, fica em posição de poder tomar o cumprimento como certo, quando pode não o ser. Com a indemnização do interesse positivo o promissário (credor) não “internaliza” o efeito sobre o promitente (devedor) de um seu eventual excesso de confiança. Antes, num tal regime de indemnização pelo interesse contratual positivo (tal como com o de execução específica), o credor, confiando em excesso, tratará um investimento como, por exemplo, a preparação para a receção da prestação ou a publicitação de um evento em que o devedor participa, como tendo um retorno certo, quando na realidade é incerto, devido à possibilidade de não cumprimento³³.

Entre os outros incentivos criados pela indemnização assumem também papel de relevo os incentivos para as partes *tomarem precauções* para evitarem ou minimizarem as possibilidades e as perdas resultantes do não cumprimento. A indemnização pelo interesse contratual positivo pode, como referimos, criar o incentivo correto para o promitente (devedor) tomar precauções contra eventualidades que aumentem o custo do cumprimento ou até o possam impossibilitar. Mas, se a indemnização for compensatória, por variar em termos marginais com os prejuízos do promissário, já este (credor), além de ter incentivos para um excesso de investimento em confiança, não terá incentivos para tomar suficientes

33 Suponha-se, por ex., que o vendedor se vincula a produzir um bem especializado para o comprador, e que este faz um investimento também específico, sem qualquer valor a não ser para aquele bem, com o objetivo de aumentar o valor do bem. Com a indemnização do interesse no cumprimento, se o devedor (vendedor) não cumprir, o credor receberá sempre a diferença entre o valor que realizaria e o preço. Isto é, o credor receberá sempre o retorno do seu investimento, mesmo que o negócio não seja cumprido e o investimento que ele fez fique na realidade desperdiçado. O rendimento marginal que o comprador tira do investimento excede o rendimento social (pois aquele é desperdiçado), conduzindo a um investimento na confiança ineficiente, por demasiado elevado.

precauções, quer para incorrer em custos, ou para induzir, ou ajudar, o devedor a evitar o não cumprimento, quer mesmo para evitar as suas próprias perdas, numa situação de típico “risco moral”.

O alargamento do âmbito da análise teórica, do cumprimento eficiente para o investimento eficiente, revela, portanto, que os benefícios potenciais da indemnização pelo interesse contratual positivo ao promover decisões eficientes *ex post*, de cumprir ou não cumprir, podem logo ser neutralizados pela falta de incentivo dessa regra para motivar a confiança e o investimento eficientes *ex ante*.

Aliás, o problema também não se pode resolver alterando a regra para a indemnização do *interesse na confiança*.

Em geral, o interesse na confiança, além dos problemas de incentivo suficiente ao cumprimento, resultará em investimentos de confiança *ainda mais ineficientes*, por *mais excessivos*, do que a indemnização do interesse no cumprimento, pois garante ao credor o pagamento dessas despesas ainda que elas não resultem em qualquer aumento do excedente contratual³⁴ – colocando o

34 Ilustremos este problema com a alteração do ex. referido. Suponhamos que *B*, que comprou a *A* cadeiras por 1 000 000, tem logo na sequência da conclusão do contrato a possibilidade de realizar investimentos de 150 000, que aumentariam o valor das cadeiras, para si, em 200 000, de 1 100 000 para 1 300 000 (por ex., viajando para negociar imediatamente a futura venda dessas cadeiras numa oportunidade única, comprando mesas ou mandando fazer almofadas que só pudessem ser utilizadas com aquelas cadeiras), mas que existe uma probabilidade de 50% de *A* não cumprir, por os seus custos de produção terem aumentado (ou por ter aparecido uma oferta de um terceiro que fosse melhor nesse montante). Numa tal situação, a indemnização pelo interesse positivo não conduz a níveis eficientes de confiança, antes resulta num excesso de investimento de confiança, ao fazer com que *B* atue como se o cumprimento fosse certo: assim, *B* sempre gastará os 150 000, pois, da sua perspectiva, ou o devedor cumprirá ou ele será indemnizado por 1 300 000. Todavia, se se vier a verificar o não cumprimento (“eficiente”) por *A*, o excedente agregado do contrato seria maior se *B* tivesse deixado de investir na confiança. Nesta hipótese de falta de investimento na confiança, o vendedor responderia por 1 100 000 menos o preço de 1 000 000 do contrato, enquanto se o comprador tiver confiado recebe 300 000 do vendedor, mas gastou 150 000, ficando melhor em 50 000 do que se não tivesse confiado. Na primeira hipótese, o comprador ganha 100 000 e o vendedor perde 100 000; na segunda, o comprador ganha 150 000, mas o vendedor perde 300 000, refletindo a diferença o facto de a despesa de 150 000 efetuada pelo comprador ser desperdiçada em caso de não cumprimento. A indemnização do interesse positivo não faz, pois, com que o comprador tome inteiramente em consideração, “internalizando”, as consequências, para o vendedor, da decisão dele mesmo, comprador, de efetuar um investimento de confiança. A medida do interesse na confiança está sujeita à mesma objeção. A decisão de investimento do comprador será, então, também efetuada como se não existisse qualquer risco, mesmo que ele esteja presente (pois o cumprimento não é eficiente nalgumas hipóteses). Mas, além disso, cria-se ainda um incentivo perverso para o comprador, em certas circunstâncias, resultante do facto de a indemnização do seu investimento na confiança estar sempre assegurada, ainda que o valor do excedente contratual não aumente com esse investimento. Suponhamos que o custo de produção das cadeiras para *A* é de 1 160 000, pelo que, aplicando a medida do interesse na confiança, ele preferirá pagar antes 150 000 (investimento de confiança de *B*) do que cumprir e sofrer um prejuízo de 160 000 (1 160 000 menos o preço de 1 000 000). O não cumprimento, com indemnização do interesse na confiança, priva o comprador *B*, em relação à indemnização do interesse positivo, de um ganho de 150 000, revelando o facto de o interesse na confiança não chegar para induzir um nível bastante de cumprimento. Porém, *B*, se tiver direito a uma indemnização pelo interesse na confiança, pode conseguir que o vendedor cumpra aumentando o seu

credor numa situação de indiferença entre a realização e a não realização de despesas de confiança, dá-lhe um “seguro” para tais despesas, mesmo que não eficientes³⁵. Se, portanto, a indemnização pelo interesse positivo fornece o incentivo correto quanto à decisão do devedor sobre o cumprimento (ou sobre as precauções para reduzir os seus custos), resulta já num excesso de confiança por parte do comprador (credor), o mesmo acontecendo, de forma agravada com a indemnização pelo interesse na confiança, a qual, além de não incentivar suficientemente ao cumprimento, aumenta ainda os problemas de excesso de confiança e permite mesmo uma utilização perversa dos “investimentos de confiança” pelo credor (uma indemnização que assegurasse a escolha do nível eficiente de confiança teria de deixar ao credor inteiramente o custo do não cumprimento resultante do desaproveitamento do excesso de confiança).

Com a compensação resultante da indemnização do interesse no cumprimento, não se consegue, pois, o incentivo eficiente na outra “margem” (confiança, precauções do credor), numa manifestação do “*paradoxo da indemnização*” resultante da incompatibilidade entre a finalidade de *compensação* e uma responsabilidade *eficiente em duas margens*³⁶. Não sendo possível encontrar uma medida de danos que resulte simultaneamente em decisões eficientes quanto ao cumprimento e quanto ao investimento com base na confiança ou em precauções para evitar o não cumprimento (ou reduzir os seus custos), pode, porém, dizer-se que a indemnização pelo interesse no cumprimento é *melhor* do que a indemnização do dano da confiança na indução de decisões de não cumprimento eficientes (dando o correto incentivo ao devedor), e *não é pior* do que esta última quanto à eficiência das decisões (do credor) de confiança e de precaução, seja o não cumprimento deliberado ou não.

investimento de confiança com uma despesa adicional, não produtiva, por ex., de 20 000, subindo o total das despesas de confiança de 150 000 para 170 000 (pelo que o vendedor cumpre). O problema do não cumprimento excessivo é, aqui, curado, mas à custa da realização de despesas de confiança desnecessárias (P. MAHONEY, *ob. e loc. cit.*).

35 Por isso se refere, por vezes, como solução para o excesso de confiança do promissário, a possibilidade de um “*antisseguro*”: as partes acordariam que, em caso de não cumprimento, o promitente pagaria a indemnização a um terceiro (a companhia de “*antisseguro*” e não ao promissário); como compensação para atribuir um tal direito à companhia de “*antisseguro*”, o promissário receberia dessa companhia antecipadamente um pagamento igual ao valor esperado desse direito. Este “*antisseguro*” visaria melhorar os incentivos das partes não as isolando totalmente do risco.

36 Tal “*paradoxo*” pode também ver-se se atentarmos na impossibilidade de eliminação do “risco moral” nas duas margens. Enquanto o direito dos seguros encontrou nas franquias, nas deduções, no co-seguro e outros mecanismos a forma de lidar com o problema “risco moral” do segurado, o regime contratual não consegue resolver o problema do “risco moral” induzido por uma indemnização por não cumprimento inteiramente compensatória: diminuir a indemnização de um credor pelo não cumprimento até ao montante da sua confiança razoável (eficiente) elimina o problema da vítima, mas fá-lo criando outro problema de *moral hazard*, sob a forma de um incentivo demasiado grande para o não cumprimento pelo devedor.

Justamente por este problema, foram propostas *soluções intermédias*, de indemnização do interesse positivo até ao montante que seria necessário se o promissário tivesse efetuado investimentos de confiança eficientes (para resolver o problema do excesso de confiança). Tratar-se-ia de limitar a indemnização ao investimento de confiança eficiente, ou, em termos menos precisos, “razoável” (em traços gerais, a doutrina da “mitigação” dos danos, ou da culpa do lesado, poderia servir para tal), deixando a partir daí o risco desse investimento ao promissário.

Um dos problemas para a eficiência do interesse positivo ou no cumprimento é, como se viu, o facto de este obrigar o inadimplente a fornecer um “seguro” à vítima contra quaisquer causas de não cumprimento. Ora, o efeito das regras indemnizatórias *sobre a afetação do risco* tem também de ser considerado: essa regra pode alterar o risco a que cada parte fica exposta, o que é importante se elas não forem *neutras em relação ao risco*. Como o interesse contratual positivo compensa o lesado pelo não cumprimento, pode ser desejável como forma de “seguro” se este for avesso ao risco, ao eliminar o grau de variabilidade da sua situação. Todavia, ao mesmo tempo, tal indemnização aumenta a variabilidade da posição do inadimplente, que pode ter de indemnizar – o que é também um risco –, e este também pode ser avesso ao risco. Tais considerações podem levar as partes a preferir (por produzir um maior nível agregado de bem-estar) uma indemnização menor, entre as medidas do interesse na confiança e do interesse no cumprimento, ou a especificar mais o contrato, quer quanto às causas de responsabilidade quer quanto ao montante da indemnização. A formulação precisa depende, porém, de fatores como os *níveis comparativos de propensão ao risco*, o tipo de risco (negativo ou positivo, monetário ou não) e a disponibilidade, ou não, de seguro por terceiro. Dada a complexidade e os custos destas informações e cálculos, bem como a possibilidade de erro, parece plausível que os tribunais tenham dificuldade em determinar montantes precisos, em particular se as partes *não especificaram* no contrato qualquer consequência, *liquidando antecipadamente* a indemnização em conformidade com a relativa propensão ou aversão ao risco.

Uma explicação completa dos incentivos que os “remédios” indemnizatórios podem representar para a conduta das partes contratuais não pode, aliás, deixar também de considerar, para ser minimamente realista, os referidos efeitos “*dinâmicos*”, em particular, logo na fase da *formação do contrato*, sobre o nível da atividade de celebração de contratos.

A análise dos “remédios” considerando não só os incentivos depois da conclusão do contrato, mas também os incentivos pré-contratuais, e em particular os incentivos à celebração do contrato, não foi logo efetuada, apesar de ser feita referência à ligação entre o nível de atividade (contratação) e os incentivos

à decisão de cumprir já desde o início da década de 70 do século xx. A escolha dessas consequências considerando também tais incentivos pré-contratuais é uma área que tem sido objeto de alguns tratamentos, mas não se encontra ainda suficientemente desenvolvida.

Uma consideração fundamental na escolha do “remédio” para o não cumprimento há de ser, na verdade, também o efeito que tal escolha terá sobre a *disponibilidade das pessoas para concluir contratos*³⁷.

Há quem defenda a regra da indemnização do interesse positivo, ou, mesmo, de medidas mais do que compensatórias, justamente sustentando a existência de uma ligação entre a decisão de cumprir ou não cumprir e os incentivos para a atividade de contratação futura. Deixando, porém, de lado uma análise baseada nos efeitos da assimetria informativa quanto à disponibilidade da outra parte para cumprir – análise que, coerentemente, levaria, sem quaisquer mecanismos compensatórios, a um colapso do mercado devido à existência de reparações não integrais, num resultado que é, porém, desmentido pela própria vitalidade económica e social da instituição que é o contrato³⁸ –, o efeito sobre

37 Note-se que, mesmo na ótica que vê a indemnização como consequência do não cumprimento, e como incentivo para evitar este, não pode deixar de considerar-se o efeito sobre a celebração do contrato. Como DAVID D. FRIEDMAN, “An Economic Analysis of Alternative Damage Rules for Breach of Contract”, *JLE*, 32, 1989, pp. 281-310 (284) salienta, “a regra indemnizatória fixa o custo de violar o contrato e assim o incentivo do comprador para o evitar. Há, porém, duas formas de evitar o não cumprimento. Uma é não deixar de cumprir o contrato e a outra é logo não o concluir. A regra do interesse no cumprimento fornece o incentivo correto apenas numa das duas margens”.

38 Nessa perspetiva, notar-se-ia que a influência sobre a contratação pode ser indireta, pois se o montante da indemnização for menor do que o interesse positivo existe um incentivo para o devedor não cumprir (ou não existirá um incentivo suficiente ao cumprimento). É claro que se as pessoas não cumprirem suficientemente os seus contratos, estes deixarão de ser celebrados na medida em que o seriam se fossem mais cumpridos, e, além disso, mesmo que a maioria das pessoas cumpra, a existência de casos de não cumprimento sancionados com indemnizações menores do que o valor integral do cumprimento (indemnizações “infracompensatórias”) pode ter como efeito a redução da disponibilidade da contraparte para pagar o mesmo montante que pagaria se o cumprimento (ou uma indemnização total deste) estivesse garantido. Ora, se o preço desce, os contratantes “honestos” (ou “diligentes”), que sempre cumpririam, deixarão de conseguir cumprir obtendo o mesmo rendimento, e tenderão a retirar-se (ou a diminuir o grau de cumprimento). O efeito desta assimetria informativa quanto à probabilidade de cumprimento é a conhecida espiral negativa de preço e não cumprimento, até o mercado colapsar e deixarem de concluir-se contratos, que ficou conhecida na economia de informação como “*lemons effect*”. Estaríamos, com efeito, perante uma aplicação do conhecido problema dos “*chaços*” (*lemons*), posto em evidência por GEORGE AKERLOF, “The Market for Lemons: Qualitative Uncertainty and the Market Mechanism”, in *Quarterly Journal of Economics*, 84, 1970, pp. 488-500. Todavia, se esta análise estivesse correta para toda a formação dos contratos, sem mais, seria difícil entender como é que a instituição do contrato sobreviveu e se expandiu. Mesmo com a indemnização do interesse no cumprimento é evidente que não foram eliminados todos os casos de não cumprimento ineficiente, nem os lesados com o não cumprimento são na prática sempre garantidos contra aquele em termos de ficarem em situação exatamente igual à que estariam se o contrato tivesse sido cumprido. Todavia, o mercado contratual não colapsa, as pessoas continuam a celebrar contratos, a mostrar que o efeito em causa não tem, na realidade, quanto à celebração de contratos, um efeito destrutivo tão relevante como parece. Tal realidade pode ter várias explicações (desde a previsão de cláusulas de liquidação antecipada da indemnização ou de cláusulas penais, quando admissíveis, até

o nível de contratação verificar-se-á através do preço, pois este será tanto maior quanto maior for a medida da indemnização a que o devedor fica obrigado em caso de não cumprimento, e influenciará diretamente a decisão sobre a contratação³⁹. A alteração da indemnização pode funcionar, assim, como uma *alteração de preço*, com efeitos sobre o número de contratos concluídos, que variam, designadamente consoante a estrutura do mercado em causa⁴⁰.

Há, além disso, outras formas pelas quais as medidas do interesse positivo e do interesse negativo (incluindo as oportunidades perdidas), em contextos de mercado onde não coincidam, influenciam diversamente incentivos pré-contra-tuais.

Assim, a indemnização do interesse positivo pode levar *os promitentes a tomar precauções* não eficientes (excessivas) ao decidir se devem ou não contratar. Em geral, as medidas indemnizatórias podem afetar os incentivos de ambas as partes *para refletir cuidadosamente sobre o contrato* (a decisão de o celebrar e o estudo do seu conteúdo) antes da sua celebração, defendendo-se que a indemnização do interesse na confiança é insuficiente para assegurar a ponderação adequada. A medida da indemnização pode ainda afetar os incentivos para as partes ponderarem *com que partes* devem contratar (e a que preço), ou, ainda, para gastarem mais tempo e recursos a *procurar partes* que poderiam estar dispostas a contratar.

A escolha dos “remédios” indemnizatórios pode, pois, influenciar *diversos outros comportamentos*, além das decisões de tomar precauções, de cumprir e de confiar. Se a teoria mais antiga se preocupou apenas com os efeitos do não cumprimento (o “não cumprimento eficiente”), e, depois, com os efeitos sobre

outras formas de “sinalizar” as qualidades da contraparte ou da prestação, passando pela previsão legal de consequências “supracompensatórias”, como, nos direitos continentais, a possibilidade de execução específica), podendo dizer-se que o problema parece suscitar-se sobretudo em certos mercados, como aqueles em que é difícil ao comprador observar antecipadamente a qualidade da prestação – por ex., o mercado dos automóveis usados, tendo sido justamente em relação a eles que G. AKERLOF salientou o referido “efeito dos ‘chaços’” (isto é, automóveis usados sem qualidade, ou *lemons*).

39 Por ex., se o credor (comprador) faz despesas de confiança no montante de 200 000, e o devedor (vendedor) não cumpre, o bem-estar do primeiro diminui em 200 000, sendo necessário que a indemnização atinja esse montante para não ficar em pior situação do que se não tivesse contratado (supondo que não existem custos de oportunidade). O problema que se põe é o de saber se o comprador está disponível para pagar mais por um contrato que preveja a indemnização, mais vantajosa para ele, no montante do interesse no cumprimento, e se as partes preferirão esse contrato a um outro, que preveja a indemnização apenas do dano da confiança.

40 Em geral, porém, afigura-se que o espectro de possíveis preços a que o contrato será celebrado, aumentando o bem-estar de ambas as partes, é maior prevendo-se a indemnização do dano da confiança do que a do dano de cumprimento: ou seja, a medida da indemnização pelo *interesse na confiança* cria um espectro de possibilidades de negociação maior, que poderia aumentar o número de contratos concluídos, pelo que seria preferível se se tivesse em mente apenas a eficiência na contratação.

o nível de confiança do destinatário, “grande parte da história da análise económica do contrato desde cerca de 1980 consistiu num aumento gradual da lista de incentivos que poderiam ser afetados pelos remédios para a violação” do contrato⁴¹, pondo em causa a afirmação de que a medida da indemnização pelo interesse no cumprimento cria os incentivos preferíveis em todas as dimensões.

Ora, mesmo sem uma identificação completa de todos esses comportamentos e da forma como eles são influenciados, pode concluir-se que a *pluralidade de decisões* em causa e dos efeitos, exige, não só que se consiga *identificar* estes, como a opção por um *critério* em função do qual se possa decidir *qual é o efeito mais importante* na situação de facto particular. Quando se decide ir além de uma avaliação da consequência corretiva adequada ao caso individual, para procurar antes criar incentivos para o comportamento das partes, franqueia-se, assim, o limiar de um universo complicado e incerto, com muitos e complexos factos, e circunstâncias que podem determinar a escolha da sanção – e isto, tanto mais quanto, pelo menos nos sistemas de direito legislado, se torna também necessária uma consideração em alguma medida *tipicizadora*, a não se querer deixar a determinação da medida do dano, sem mais indicações, sempre ao tribunal.

Mais recentemente, têm vindo a ser propostas soluções indemnizatórias “sofisticadas”, quer estendendo a análise a novas possíveis cláusulas contratuais quer facultando a otimização tanto das decisões *ex post* sobre o cumprimento como dos incentivos *ex ante*, para a contratação e a realização de investimentos específicos de confiança, mediante o desenho de mecanismos aptos a induzir certos resultados.

A complexidade é ainda aumentada – como referimos – pela necessária incorporação na análise da possibilidade de as partes *renegociarem* os termos do seu acordo.

Além da atenção ao efeito sobre as negociações pré-contratuais, é, na verdade, necessário analisar os efeitos, para a escolha de “remédios” eficientes, da possibilidade de renegociação e dos custos desta (podendo também ser importante saber quais as condições sob as quais o custo de previsão de uma medida de indemnização eficiente, ou do afastamento de uma medida ineficiente no momento da celebração do contrato, é maior ou menor do que o custo da renegociação no momento do cumprimento). O contributo da medida indemnizatória como incentivo ao cumprimento será menor se as partes contratantes puderem renegociar os seus acordos, tendo como resultado sempre

41 Assim, RICHARD CRASWELL, “Two Economic Theories of Enforcing Promises”, in PETER BENSON (org.), *The Theory of Contract Law*, cit., 2002, pp. 19-44 (26-8).

uma decisão eficiente sobre o cumprimento. A análise tradicional aceitava que a renegociação no momento do cumprimento estaria sujeita a custos proibitivos. Mas se, pelo contrário, a renegociação não tiver *quaisquer custos*, a regra indenizatória é, como dissemos, *irrelevante* em termos de eficiência, pois as partes *negociarão sempre* até obter as decisões eficientes sobre o cumprimento e o não cumprimento e o investimento na confiança (como resulta do chamado “teorema de Coase”).

De todo o modo, parece certo que em muitas circunstâncias o contrato na realidade *não será renegociado*, desde logo, por as partes não estarem em contacto entre si quando surgem dificuldades e uma das partes necessita de tomar uma decisão rapidamente (por exemplo, no decurso de um processo produtivo, precisando de decidir se prossegue com ele ou não, ou perante uma nova oferta que expira logo, de decidir se a aceita ou não), ou por existirem entre as partes assimetrias informativas que podem comprometer o resultado da negociação. Nesta situação mais plausível, em que a renegociação tem *custos* (de transação), mas nem sempre proibitivos, a escolha do “remédio” é mais difícil, e exige a consideração desses custos, sendo que a atribuição inicial do direito ao (não) cumprimento pode então (contrariamente ao “teorema de Coase”) ter efeitos relevantes sob o ponto de vista da eficiência.

E mesmo aceitando que a renegociação tende a verificar-se, parece também plausível que nas condições reais nem sempre resulte num resultado eficiente. Muitas vezes, a renegociação permite às partes conseguir a maximização do excedente agregado da transação, evitando decisões de não cumprimento ineficientes. Assim, por exemplo, se a consequência do não cumprimento for uma indemnização maior do que o dano de cumprimento, ou se for a execução específica, o devedor que seria levado a cumprir mesmo que o custo, para ele, excedesse o valor do cumprimento para o credor, pode, para evitar tal resultado ineficiente, propor ao credor efetuar-lhe um pagamento *a latere*, para que este o exonere da sua obrigação, e sem que, portanto, a medida da indemnização seja indispensável para atingir um resultado eficiente e sem que tenha de estar antecipadamente prevista no contrato.

Noutras situações, porém, o resultado da renegociação pode ser apenas um *sucedâneo imperfeito* para uma *cláusula que liquide antecipadamente* a indemnização ou para uma *medida adequada* de indemnização legalmente prevista, quer porque a renegociação apenas pode afetar decisões futuras sobre o cumprimento, e não a conduta que já se verificou antes dela, quer por causa dos riscos que são afetados ao potencial inadimplente (por exemplo, o risco de vir a ter de suportar um custo de produção muito elevado), para o caso de falhanço da renegociação (e que seriam limitados com uma liquidação convencional ou uma previsão legislativa antecipada do montante da indemnização).

Por outro lado, a presunção de que a possibilidade de renegociação é desejável para ambas as partes, por lhes permitir melhorar a sua situação quando surgem dificuldades e por os dispensar de aumentar a complexidade do clausulado contratual, nem sempre pode considerar-se procedente, pois a possibilidade de renegociação tem também efeitos negativos.

A perspetiva da possibilidade de renegociação afeta, também, os incentivos das partes para investir na relação contratual (como veremos mais pormenorizadamente a seguir, a propósito da análise económica da responsabilidade pré-contratual). Quem realiza um investimento confiando na relação contratual tenderá a antecipar, nesse caso, a captura pela contraparte, mediante renegociação (de pelo menos parte), do excedente criado por esse investimento. Tal situação de possível *hold-up* afeta o estabelecimento de incentivos à realização de investimentos específicos de confiança a um nível eficiente.

Além disso, a possibilidade de renegociar pode funcionar em detrimento das partes por evitar que elas se comprometam firmemente a certos resultados logo no contrato inicial, podendo ser do interesse de ambas a criação de obstáculos jurídicos ou práticos (como a exigência de várias aprovações intermédias, ou a intervenção de terceiros) à renegociação ou a determinação antecipada do resultado desta (da afetação do excedente adicionalmente criado).

Tem, assim, sido salientada a conveniência da passagem de uma visão *dualista* da decisão do devedor (cumprir ou não cumprir) para uma perspetiva *triádica*, em que o devedor opta entre *cumprir*, *não cumprir* ou *renegociar* a “saída” do contrato. Em muitos contextos, um dos objetivos principais do acordo inicial será justamente condicionar, ou fixar “pontos de ameaça”, adequados para uma renegociação subsequente. Objetivo cuja prossecução, porém, remete para a alternativa, com que as partes estarão normalmente confrontadas, entre, por um lado, a celebração de contratos *mais completos* (com os correspondentes custos) que facilitam a indução tanto da contratação e cumprimento eficiente como do investimento eficiente, mas apenas podem resultar se não existem perspetivas de renegociação, ou, por outro lado, a conclusão de contratos *mais simples*, que as partes podem renegociar para conseguir negócios eficientes sem comprometer a eficiência do investimento.

No final de contas, parece, pois, difícil fornecer incentivos – em soluções necessariamente não muito “sofisticadas”, pois a própria complexidade pode dificultar os cálculos dos agentes e, porventura, acarretar outros custos e efeitos – *em todas as dimensões relevantes* dos negócios em causa. O não cumprimento eficiente, a confiança eficiente, a precaução e redução de danos eficiente, a afetação de risco eficiente, a eficiência na ponderação pré-contratual e na escolha da contraparte, o grau e o resultado adequado da renegociação – não parece que todos estes múltiplos objetivos possam ser *simultaneamente*

induzidos ou *promovidos* por uma *única regra*, sem qualquer *compromisso*, sendo então necessário apurar em que se traduz este e qual a importância relativa dos incentivos preferidos ou preteridos⁴². É preciso, pois, definir prioridades para escolher a dimensão em que é mais importante atuar, fornecendo os incentivos corretos.

E, sobretudo, o que importa reiterar é que, como se tem salientado na doutrina, não existe qualquer razão para supor que a totalidade dos efeitos económicos apontará sempre para uma *certa medida, definida* como o interesse no cumprimento ou o interesse na confiança (ou o “interesse na restituição”). Pelo que, na perspetiva do “desenho” dos incentivos corretos para a conduta das partes, não existe razão para supor que as situações correspondentes integralmente ao cumprimento, à conclusão de um contrato alternativo, ou à não conclusão do contrato, apresentem alguma pretensão especial (em lugar, por exemplo, de uma certa percentagem dos correspondentes “interesses”) para se afirmarem como linhas ou *pontos de referência* decisivas para a medida do dano. Dever-se-ia, assim, rejeitar a própria *utilização dos conceitos* de interesse na confiança e interesse no cumprimento como “*fasquias fixas*”, conceptualmente demarcadas, reduzindo-se, antes, numa atitude de “despreocupação semântica”, a meros “*pontos focais*”, que apenas facilitam supletivamente a formação de um conteúdo obrigacional⁴³.

Ora, não se pode deixar de observar que estas conclusões, perfeitamente coerentes na lógica económica, parecem comprometer irremediavelmente *mesmo a pretensão descritiva* da análise económica dos “remédios” indemnizatórios – uma das “joias da coroa” do instrumentalismo efficientista na teoria dos contratos.

Não se nega a obtenção de alguns importantes resultados da abordagem económica, ao pôr em relevo a ligação entre os diversos “remédios” e os incentivos das partes em diversas dimensões. Mas, apesar de tais contributos, a explicação instrumentalista dos “remédios” contratuais (sobretudo com os refinamentos introduzidos à doutrina do “não cumprimento eficiente”) é insuficiente, estando longe das *soluções do direito vigente*, em particular quanto à preferência clara pelo interesse positivo, com atribuição ao credor do lucro que para ele resultaria do contrato. Designadamente, em várias hipóteses tal atribui-

42 Outro aspeto pouco tratado, que igualmente aumenta a complexidade do problema, é a circunstância de algumas consequências jurídicas poderem ser cumulativas, e não alternativas, tendo a parte a possibilidade de optar por cumular aqueles “remédios” que pretende exercer.

43 V. RICHARD CRASSWELL, “Against Fuller and Perdue”, 2000, pp. 107-11, e “How We Got This Way: Further Thoughts on Fuller and Perdue”, *ILS, Symposium: Fuller and Perdue*, 2001, artigo 2 (www.bepress.com/ils/iss1/art2), *passim*. Em Portugal, FERNANDO ARAÚJO, “Da tutela negativa da confiança ao ‘paradoxo da indemnização’”. Uma análise económica dos contratos”, in *Estudos em memória do Professor Doutor Marques dos Santos*, Coimbra, Almedina, 2005, pp. 441-592 (556, ss., 558, 561).

ção de uma indemnização pelo interesse positivo seria contrariada pelo efeito de excesso de confiança do promissário, por não induzir precauções suficientes, ou por não ter em conta a relativa aversão das partes ao risco. E sobretudo, de forma decisiva, o instrumentalismo não tem explicação para a *insistência* do direito dos contratos nas medidas do interesse positivo e do interesse na confiança como *categorias formais* – o que não significa necessariamente que tenham existência *a se*, a “gravitar fantasmagoricamente” em torno do regime legal⁴⁴ –, vendo-se forçado a remeter tal insistência (ou a sua subsistência) para uma espécie de *resistência “ideológica”*⁴⁵.

Esta incapacidade de explicar a insistência do direito positivo na atribuição ao credor de uma indemnização que, seja qual for o seu montante (e salvo correções pontuais, como a que poderá resultar do artigo 494.º), parta da correspondência ao dano de cumprimento (ao interesse no cumprimento ou interesse positivo), que garante ao credor o lucro do negócio, denuncia, a nosso ver, a *limitada percepção do fundamento da solução jurídica* em causa. Como Craswell observou, mesmo quando a abordagem económica possa recomendar a indemnização do interesse positivo como eficiente, trata-se de mera coincidência, e a categoria do interesse no cumprimento, ou do interesse positivo, não terá desempenhado qualquer papel relevante na fundamentação⁴⁶.

Mas, justamente, o direito positivo, ao remeter para o critério da causalidade do prejuízo pelo não cumprimento (artigo 798.º do Código Civil Português) remete para uma tal indemnização pelo interesse positivo, a qual é aceite na doutrina e na jurisprudência. A ordem jurídica, ao prever a indemnização pelo dano de cumprimento, remete, pois, em geral para um *efeito do negócio* (da promessa, ou do contrato) que não é descrito ou explicado em termos instrumentais, apenas como um esforço de promover a eficiência (o que, porém, também não significa que deva ser explicado em termos meramente formais, podendo servir outros valores substanciais)⁴⁷.

44 F. ARAÚJO, “Da tutela negativa da confiança...”, cit., p. 558 – é justamente na explicação de tais regimes, na sua adequação a eles (ou, se se quiser, na explicação de tal pretensa “gravitação fantasmagórica”) que consiste a missão da dogmática jurídica, contribuindo para a racionalização dessas soluções, para uma adequada delimitação (ou extensão) e para a sua fundamentação valorativa.

45 F. ARAÚJO, “Da tutela negativa da confiança...”, cit., p. 562.

46 R. CRASWELL, “Against Fuller and Perdue”, cit., p. 107.

47 A tentativa de racionalizar em termos instrumentais tal solução, ligando-a com a eficiência do ‘valor de garantia’ do lucro do negócio não pode deixar de ser entendida justamente como tal – isto é, como mera racionalização, produto de um compromisso com uma explicação em termos instrumentais que precede o compromisso com o direito, e não se dirige a este último no seu próprio terreno.

Intellectual property rights: infringement and calculation of damages

Henrique Sousa Antunes

Professor of Law

Católica Lisbon Law School, Universidade Católica Portuguesa

A Comment on the Judgment of the Court of Justice of the European Union (Fifth Chamber) *Stowarzyszenie 'Olawska Telewizja Kablowa' v Stowarzyszenie Filmowców Polskich*, 25 January 2017 (Case C-367/15)

1. Background to the Comment

In the Judgment *Stowarzyszenie 'Olawska Telewizja Kablowa' v Stowarzyszenie Filmowców Polskich*, of 25 January 2017 (Case C-367/15), the Court of Justice pronounced on the compatibility of national legislation granting the rightholder whose economic rights of copyright have been infringed the possibility of, as an alternative to compensation for the damage actually suffered, requesting payment of a sum which corresponds to twice the fee that would be due for the permission to use the work, with Article 13 of Directive 2004/48/EC of the European Parliament and of the Council, of 29 April 2004, on the enforcement of intellectual property rights¹.

This was a request for a preliminary ruling presented by the Polish Supreme Court which, considering the application of the national legislation in view of Article 13(1) (b) of Directive 2004/48, questioned the legitimacy of the injured party requesting compensation corresponding to twice the appropriate fee, without being required to prove the loss and the causal relationship between the event and such loss, as an alternative to redress for the damage suffered, calculated in line with general principles. The Court also asked whether an order to pay pre-established damages via the application of a multiple of the appropriate fee would be permissible under Article 13 of the Directive, given that recital 26 of the preamble to the Directive excludes the introduction of punitive damages from the Directive's aim.

This ruling of the Court of Justice is significant for three key reasons: 1) it recognises that the Directive seeks minimal harmonisation of the rules on compensation for infringement of intellectual property rights; 2) it separates the pre-established compensation from proof of the specific extent of the damage; 3) it rejects the argument that the Directive prohibits the fixing of punitive damages. It should be stressed, however, that, unjustifiably, it silences the compatibility of such damages with the Directive. It will be of interest to analyse each of the aforementioned points.

2. The minimal level of protection

Article 2(1) of Directive 2004/48 lays down: "Without prejudice to the means which are or may be provided for in Community or national legislation, in so far as those means may be more favourable for rightholders, the measures, pro-

¹ The Directive is published in the "Official Journal of the European Union", L 195, 2.6.2004, pp. 16 to 25.

cedures and remedies provided for by this Directive shall apply, in accordance with Article 3, to any infringement of intellectual property rights as provided for by Community law and/or by the national law of the Member State concerned.”

In this case C-367/15, the Advocate General Eleanor Sharpston, in the Conclusions presented on 24 November 2016, excluded the applicability of the minimal harmonisation clause to compensation for loss and damage (Article 13 of the Directive), drawing a distinction between *means* and *measures*. In short, the *measures* that the Directive included must be interpreted in order to ensure equivalent protection in the different Member States. A different conclusion is only possible regarding *means* which, with a view to the same aim pursued by the Directive, are not ruled by it (paragraphs 49 to 54). She writes: “53. [...] I do not see that Article 2(1) can serve as a basis for an argument which states that Directive 2004/48 introduces only minimal harmonisation in those areas which it does cover. First, such a proposition runs counter to the wording of that provision, which refers not to ‘measures’ which are or may be provided for in EU or national legislation but to ‘means’. Second, it fails to reflect the general scheme of the directive, which, as indicated in recital 8, is to ensure that intellectual property rights enjoy an equivalent level of protection throughout the European Union. In so far as the directive lays down rules applying to a particular form of remedy, as it does in the case of damages, the rules should in my view be the same throughout the European Union.”

This statement presupposes, however, that, if the national legislation includes a measure of a punitive nature in the context of compensation, this must be classified as compensation and, thus, its legitimacy in the light of Directive 2004/48 is excluded. In fact, when pronouncing on the claims of the Polish government that the Court of Justice should have regard to the Judgment rendered by it in the case of *Maria Auxiliadora Arjona Camacho v Securitas Seguridad España, SA*, of 17 December 2015 (Case C-407/14), according to which the provision of the payment of punitive damages to whomever has been a victim of discrimination based on gender constitutes a proportionate and, as such, legitimate sanction, the Advocate General states: “46. I can draw no useful guidance from the judgment in question. The provision which it refers to concerns, inter alia, the payment of sums by way of penalties and not as damages. In the former context, it is natural that the calculation concerned may bear no necessary relation to the harm suffered by the victim. The same cannot be said of an award of damages which, by virtue of Article 13(1) of Directive 2004/48, must be appropriate to the *actual* prejudice suffered by the rightholder. In that context, the test of proportionality presupposes, in my view, that there be some relationship between the loss suffered and the amount claimed. I suggest that an award of punitive damages will not, by definition, satisfy that test.”

This understanding does not seem to us to be the correct one. The classification of a *means* or of a *measure* must be given by the purpose of the national regulation and not by the name it has been given. A sanction which exceeds compensation of the damage and, therefore, favours the injured party, must be understood as a private penalty and, as such, its legitimacy must be assessed. In line with this reasoning, and even if the interpretation of Article 2(1) of Directive 2004/48 was that proposed by the Advocate General, the fact that the European legislator established a regime for damages is not sufficient to restrict the assessment of the applicability of the national rule on the duty to compensate to Article 13 of the Directive.

It is, therefore, fitting to assess whether the private penalty finds support in any of the other provisions of Directive 2004/48 which, in the absence of complementary regulations, validate the means to which the national legislator resorts for such purpose. It may be noted that, in *Arjona Camacho*, interpreting Directive 2006/54/EC of the European Parliament and of the Council, of 5 July 2006, on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast), the Court of Justice distinguished between the aim of the compensation provided for in Article 18, the function of which is full compensation for the loss, and the scope of the duty to indemnify². On the latter, it considered that Article 25, on the provision of penalties, allows the national law to establish the obligation to pay punitive damages to a victim of discrimination on the grounds of sex. In other words, the European legislator's option to regulate compensation in the classic manner, limited by the extent of the damage, cohabited with the inclusion of the duty to indemnify being used to dissuade discriminatory conduct.

A different understanding cannot be drawn from Directive 2004/48. In particular, Article 3 lays down: "1. Member States shall provide for the measures, procedures and remedies necessary to ensure the enforcement of the intellectual property rights covered by this Directive. Those measures, procedures and remedies shall be fair and equitable and shall not be unnecessarily complicated or costly, or entail unreasonable time-limits or unwarranted delays; 2. Those measures, procedures and remedies shall also be effective, proportionate and dissuasive and shall be applied in such a manner as to avoid the creation of barriers to legitimate trade and to provide for safeguards against their abuse." It is true that, in line with recital 26, the aim of the rules on loss and damage was not to introduce "an obligation to provide for punitive damages but to allow for compensation based on an objective criterion while taking account of the expenses incurred by the rightholder, such as the costs of identification and research".

² The Directive is published in the "Official Journal of the European Union", L 204, 26.7.2006, pp. 23 to 36.

It is, however, also true that, as we shall see, the Directive does not rule against the legitimacy of punitive damages and that, according to recital 3, “without effective means of enforcing intellectual property rights, innovation and creativity are discouraged and investment diminished”. Recital 3 also states that it is “necessary to ensure that the substantive law on intellectual property, which is nowadays largely part of the *acquis communautaire*, is applied effectively in the Community. In this respect, the means of enforcing intellectual property rights are of paramount importance for the success of the internal market”.

It is within this scope that Article 16 of Directive 2004/48 finds justification: “Without prejudice to the civil and administrative measures, procedures and remedies laid down by this Directive, Member States may apply other appropriate sanctions in cases where intellectual property rights have been infringed.” This rule is compatible with the aim of the Directive: “[...] to approximate legislative systems so as to ensure a high, equivalent and homogeneous level of protection in the internal market” (recital 10). There are, therefore, grounds for national legislators to use their compensation rules as means of protecting intellectual property rights, despite the directive’s inclination towards the classic instruments: “[...] in addition to the civil and administrative measures, procedures and remedies provided for under this Directive, criminal sanctions also constitute, in appropriate cases, a means of ensuring the enforcement of intellectual property rights” (recital 28).

Where European law safeguards the freedom of the national legislator to fix sanctions which are effective, proportionate and dissuasive, the civil law may operate as an instrument capable of satisfying the aims sought by the European regulations. It is a known fact that the European legislative and judicial bodies have emitted contradictory indications regarding the use of punitive damages³. We may safely state, however, that the provision of a regime for compensation for resulting harm and lost profits does not preclude, in itself, recourse to punitive damages as an additional sanctioning instrument in the national law.

In this sense, the Court of Justice, in the decision being annotated, rejected the interpretation of the Advocate General, which, as we have seen, restricted the competence of the national legislator, subordinating the concept of means to the need to guarantee equivalent protection of intellectual property in the terms in which the measures were included in the Directive.

According to Eleanor Sharpston, the compensation could not have a reach other than that covered by Article 13, without prejudice to the means which the national legislation lays down for such purpose. Yet, in the Judgment we read: “21. It should be noted, first of all, that, as is apparent from recital 3, Directive

³ See Bernhard A. Koch, *Punitive Damages in European Law*, in “Punitive Damages: Common Law and Civil Law Perspectives” (Helmut Koziol/Vanessa Wilcox – eds.), Wien/New York, 2009, pp. 197 et seq.

2004/48 seeks to ensure that the substantive law on intellectual property is applied effectively in the European Union. Thus, Article 3(2) of the directive requires the measures, procedures and remedies provided for by the Member States to be effective, proportionate and dissuasive. 22. Whilst recital 10 of Directive 2004/48 refers, in this context, to the objective of ensuring a high, equivalent and ‘homogeneous’ level of protection, of intellectual property in the internal market, the fact remains that, as is apparent from Article 2(1), the directive applies without prejudice to the means which are or may be provided for, in particular, in national legislation, in so far as those means may be more favourable for rightholders. *It is quite clear from recital 7 of the directive that the term ‘means’ that is used is general in nature, encompassing the calculation of damages.* 23. Consequently, as the Court has already held, Directive 2004/48 lays down a minimum standard concerning the enforcement of intellectual property rights and does not prevent the Member States from laying down measures that are more protective (see judgment of 9 June 2016, *Hansson*, C-481/14, EU:C:2016:419, paragraphs 36 and 40)” (our italics).

Accordingly, the Court of Justice accepts that the concept of compensation can be tailored to different regimes and that the framing of Article 13 does not preclude the intervention of the national legislation in defining how to calculate the compensation. Thus, if the compensation exceeds the extent of the damage, as was set out by the aforementioned article, its legitimacy is only assessed using the criteria of Article 3(2) of Directive 2004/48. The measures, procedures and remedies adopted by Member States to guarantee enforcement of intellectual property rights “[...] shall [...] be effective, proportionate and dissuasive [...]”. This is, therefore, a case of minimal harmonisation.

The same solution can be found in other international instruments by which the States are bound, as referred to by the Court of Justice: “24. [...] in accordance with recitals 5 and 6 and Article 2(3)(b) of Directive 2004/48, it is necessary, for the purpose of interpreting the directive’s provisions, to take account of obligations on the Member States resulting from international agreements, including the TRIPS Agreement, the Berne Convention and the Rome Convention, which could apply to the dispute in the main proceedings. Article 1 of the TRIPS Agreement as well as Article 19 of the Berne Convention and Article 2 of the Rome Convention permit Contracting States to grant the holders of the rights concerned wider protection than that respectively laid down by those instruments.”

Considering all of the foregoing, the masking of the debate as to whether punitive damages are congruent with Article 13 leaves a bitter aftertaste. The Court states: “29. [...] without there being any need to rule on whether or not the introduction of ‘punitive’ damages would be contrary to Article 13 of Directive 2004/48, it is not evident that the provision applicable in the main proceedings

entails an obligation to pay such damages.” The Court takes refuge, comfortably, in giving an indemnifying aim to twice the amount of the hypothetical fee, as if that classification could avoid the lack of an exact equivalence between the amount provided for and the damage that the injured party, in fact, suffered. This is merely a word game, to hide that which is visible to all: the legitimacy of damages being used to dissuade conduct, provided that the penalty is proportionate to the harm suffered. We shall return to the topic.

3. Pre-established damages

Article 13 of Directive 2004/48 provides as regards compensation for loss and damage: “1. Member States shall ensure that the competent judicial authorities, on application of the injured party, order the infringer who knowingly, or with reasonable grounds to know, engaged in an infringing activity, to pay the rightholder damages appropriate to the actual prejudice suffered by him/her as a result of the infringement. When the judicial authorities set the damages: a) they shall take into account all appropriate aspects, such as the negative economic consequences, including lost profits, which the injured party has suffered, any unfair profits made by the infringer and, in appropriate cases, elements other than economic factors, such as the moral prejudice caused to the rightholder by the infringement; or b) as an alternative to (a), they may, in appropriate cases, set the damages as a lump sum on the basis of elements such as at least the amount of royalties or fees which would have been due if the infringer had requested authorisation to use the intellectual property right in question. 2. Where the infringer did not knowingly, or with reasonable grounds to know, engage in infringing activity, Member States may lay down that the judicial authorities may order the recovery of profits or the payment of damages, which may be pre-established.”

According to the Advocate General in this case, the application of Article 13(1) (b) constitutes a regime of exceptional application and presupposes that the injured party demonstrates a lack of disproportionality between the loss actually suffered and the compensation calculated thus. Let us see:

a) The exceptional application. Without prejudice to the rule’s silence on this, recital 26 of the Directive appears to support the interpretation that the provision is afforded a subsidiary nature, only applicable when it proves difficult to determine damages in accordance with sub-paragraph a). In view of such difficulty, the injured party has an alternative tool available to him to compensate the harm. Eleanor Sharpston writes: “37. [...] the legislature thus intended to enable the rightholder to avoid the potentially large amount of time and expense that he

might otherwise have to incur as a condition of bringing proceedings against an infringer.”

The line of reasoning followed by the Advocate General imposes on the injured party seeking to use the pre-established damages criterion the burden of proving “that it is accordingly ‘appropriate’ to do so. It must, as a minimum, either be ‘difficult to determine the amount of the actual prejudice suffered’ or there must be grounds which establish that the award of damages limited to a sum calculated by reference to subparagraph (a) is manifestly unjust or unreasonable” (38.). This understanding is not contested by the ruling of the Court of Justice;

b) The causal link between the compensation and the harm actually suffered. Eleanor Sharpston considers that the lump sum method, since it cannot be dissociated from that set out in paragraph 1 of Article 13, requires the injured party to demonstrate that the measure is not disproportionate to the damage caused: “47. [...] What the rightholder must establish is that the loss suffered and the amount claimed are not out of proportion to one another. To that degree, therefore, he must establish a causal link between the two. By the nature of the remedy provided for in subparagraph (b) of Article 13(1), it is not necessary for the rightholder to establish that relationship with mathematical certainty, since the whole point underlying that provision is to cater for circumstances in which it may be difficult or impossible to do so. But it is in my view necessary that he demonstrate some relationship and that he should not be given an entitlement to damages that are in no way commensurate to his actual loss.”

The proof required appears to us to lack grounds, for two key reasons:

1 – the demonstration, albeit approximate, of the damage that was, in fact, suffered, cancels out the effects intended by the provision of this formula for calculating, namely “while taking account of the expenses incurred by the rightholder, such as the costs of identification and research” (recital 26 of the Directive). As, indeed, the Advocate General acknowledges, writing with regard to the contribution of a system based on royalties or fees for the prevention of those charges: “37. [...] Were the position to be otherwise, there would be a risk that the requirement under Article 3(1) of Directive 2004/48 that remedies must not be unnecessarily complicated or costly or entail unwarranted delays could not be satisfied. They would not, in other words, be ‘effective’ and hence ‘dissuasive’ for the purposes of Article 3(2) of the directive.”

In this sense, the Court decided that the proof of damage is contrary to the nature of the compensation set out in Article 13(1) (b) of the Directive: “32. [...] as regards the argument that, inasmuch as the injured party could calculate the damages on the basis of twice the amount of the hypothetical royalty, he would no longer have to prove the causal link between the event giving rise to the cop-

right infringement and the loss suffered, it must be stated that that argument is based on an excessively strict interpretation of the concept of ‘causality’, under which the holder of the infringed right should establish a causal link between that event and not only the loss suffered but also its precise amount. *Such an interpretation is irreconcilable with the very idea of setting damages as a lump sum and, therefore, with Article 13(1)(b) of Directive 2004/48, which permits that type of compensation”* (our italics).

It may be added that the Court accepts the idea of adequacy between the compensation and the damage, but, so as to correspond to the cited option, places the burden on the injuring party to prove the lack of proportionality between the extent of the compensation and the injured party’s losses: “31. It is admittedly possible that, in exceptional cases, payment for a loss calculated on the basis of twice the amount of the hypothetical royalty will exceed the loss actually suffered so clearly and substantially that a claim to that effect could constitute an abuse of rights, prohibited by Article 3(2) of Directive 2004/48”;

2 – the assessment of proportionality cannot be made only regarding the extent of the damage, if we accept that the compensation performs a punitive function. In this context, the level of fault of the perpetrator or repetition of his conduct is particularly relevant. With the aforementioned reach, as will be stated below, Portuguese law makes adequate provision.

4. The absence of a prohibition on punitive damages

The Judgment states emphatically: “28. Contrary to the view that the referring court appears to take, the fact that Directive 2004/48 does not entail an obligation on the Member States to provide for ‘punitive’ damages cannot be interpreted as a prohibition on introducing such a measure.”

It is true that the Court silenced any statement in favour of the legitimacy of that measure. The legitimacy of punitive damages seems, however, to be acceptable, in general terms. This has already been stated in relation to the enshrining of a minimum level of protection.

European law bases the exclusion of punitive damages on the legislator expressly providing that it is not legitimate, a statement that is absent in Directive 2004/48, or on the disproportionality which, in the understanding of the case law, characterises such penalties.

Let us begin with the argument based on the legal prohibition. By way of an example, the possibility of an order awarding punitive damages was denied in the area of collective protection of rights and competition law:

a) In the Commission Recommendation of 11 June 2013, on common principles for injunctive and compensatory collective redress mechanisms in the Member States concerning violations of rights granted under Union Law, we may read in principle 31 (prohibition of punitive damages): “The compensation awarded to natural or legal persons harmed in a mass harm situation should not exceed the compensation that would have been awarded, if the claim had been pursued by means of individual actions. In particular, punitive damages, leading to overcompensation in favour of the claimant party of the damage suffered, should be prohibited.” The stated intention is “to avoid the development of an abusive litigation culture in mass harm situations” (recital 15)⁴;

b) In Directive 2014/104/EU of the European Parliament and of the Council, of 26 November 2014, on certain rules governing actions for damages under national law for infringements of the competition law provisions of the Member States and of the European Union, Article 3(3) provides that “full compensation under this Directive shall not lead to overcompensation, whether by means of punitive, multiple or other types of damages”⁵.

Considering the provisions cited, included in regimes where dissuasion of conduct is particularly significant, there would be grounds to consider that punitive damages are contrary to the public order of the European Union. And, yet, a different conclusion can be drawn, namely, from Regulation (EC) No 864/2007 of the European Parliament and of the Council, of 11 July 2007, on the law applicable to non-contractual obligations (Rome II)⁶. While in the initial text such opposition was stated, the final version did not adopt the solution, with Article 26, on public policy of the forum, providing a safeguard for States with regard to damages of punitive character. Recital 32 of the Regulation states that reasons “of public interest justify giving the courts of the Member States the possibility, *in exceptional circumstances*, of applying exceptions based on public policy and overriding mandatory provisions. In particular, the application of a provision of the law designated by this Regulation which would have the effect of causing *non-compensatory exemplary or punitive damages of an excessive nature to be awarded may, depending on the circumstances of the case* and the legal order of the Member State of the court seised, be regarded as being contrary to the public policy (*ordre public*) of the forum” (our italics). It is significant to note that the formula chosen subjects the inapplicability of a law which legitimises the awarding of punitive damages to the excessive reach of damages, thereby suggesting the acceptance of different effects of civil liability.

4 In the “Official Journal of the European Union”, L 201, 26.7.2013, pp. 60 to 65.

5 In the “Official Journal of the European Union”, L 349, 5.12.2014, pp. 1 to 19.

6 In the “Official Journal of the European Union”, L 199, 31.7.2007, pp. 40 to 49.

In tune with this understanding, and in the absence of a prohibition on punitive damages, it seems appropriate that the European case law limits itself to assessing the proportionality between the measure and the harm, respecting the principles of equivalence and effectiveness. Thus, the Judgment of the Court of Justice in *Vincenzo Manfredi v Lloyd Adriatico Assicurazioni SpA*, of 13 July 2006 (in Joined Cases C-295/04 to C-298/04) states: “99. [...] in accordance with the principle of equivalence, if it is possible to award specific damages, such as exemplary or punitive damages, in domestic actions similar to actions founded on the Community competition rules, it must also be possible to award such damages in actions founded on Community rules. However, Community law does not prevent national courts from taking steps to ensure that the protection of the rights guaranteed by Community law does not entail the unjust enrichment of those who enjoy them.”

It cannot be denied, however, that there is a temptation to recognise punitive damages as a measure which is, by its nature, disproportionate and, thus illegitimate. According to European legislation, applicable enforcement instruments must be effective, proportionate and dissuasive. Regarding the case being discussed here, the Advocate General writes: “46. [...] In the former context [payment of sums by way of penalties], it is natural that the calculation concerned may bear no necessary relation to the harm suffered by the victim. The same cannot be said of an award of damages which, by virtue of Article 13(1) of Directive 2004/48, must be appropriate to the *actual* prejudice suffered by the right-holder. In that context, the test of proportionality presupposes, in my view, that there be some relationship between the loss suffered and the amount claimed. *I suggest that an award of punitive damages will not, by definition, satisfy that test*” (our italics).

Even if Members States were forbidden from adopting measures which are more generous to the injured party, which, as we have seen, is not accepted, Eleanor Sharpston’s interpretation seems to forget that, in Article 13, the legislator uses parameters other than equivalence between the compensation and the loss. In this context, we may refer, immediately, to the distinction between the rules on culpable infringement and on bona fide infringement. Other examples are the inclusion of unfair profits obtained by the infringer among the criteria for setting the damages, the provision of the alternative method as a minimum compensation due to the injured party, accepting, in these terms, the assessment according to fairness, or receipt of compensation for moral damages. It is, therefore, accepted that the fault of the perpetrator plays a significant role in Article 13.

The reference in that article to appropriate damages seeks only to guarantee that the injured party obtains redress for the damage suffered, without reduc-

ing proportionality to a mathematical operation of equivalence. On the concept of appropriateness, we can read, for example, in the Judgment of the Court of Justice in *Nadine Paquay v Société d'architectes Hoet + Minne SPRL*, of 11 October 2007 (Case C-460/06): “45. [...] the objective is to arrive at real equality of opportunity and cannot therefore be attained in the absence of measures appropriate to restore such equality when it has not been observed. [...] 46. [...] Where financial compensation is the measure adopted in order to achieve the objective previously indicated, *it must be adequate, in that it must enable the loss and damage actually sustained as a result of the discriminatory dismissal to be made good in full in accordance with the applicable national rules [...]*” (our italics).

With this aim, the Court of Justice, in the Judgment in *Christian Liffers v Producciones Mandarina SL and Mediaset España Comunicación SA*, of 17 March 2016 (Case C-99/15), decided that the silence of the European legislator was not counter to the possibility of the injured party combining the claim for compensation calculated on the basis of the hypothetical fee with compensation for non-material damage suffered: “26. [...] setting the amount of damages due as a lump sum on the basis of hypothetical royalties alone covers only the material damage suffered by the intellectual property rightholder concerned; consequently, for the purposes of providing compensation in full, that rightholder must be able to seek, in addition to the damages thus calculated, compensation for any moral prejudice which he has suffered.”

The recognition that Directive 2004/48 establishes a minimal harmonisation facilitates acceptance of punitive damages. It will be recalled that the Court of Justice safeguarded the legitimacy of the Member States to provide for compensation which is more favourable to the injured party (20. to 24.): “25. Accordingly, Article 13(1)(b) of Directive 2004/48 must be interpreted as not precluding national legislation, such as that at issue in the main proceedings, which provides that the holder of economic rights of copyright that have been infringed may require the person who has infringed those rights to compensate for the loss caused by payment of a sum corresponding to twice the amount of a hypothetical royalty.” Regarding the intervention of national rights, we may read in recital 17 of the Directive: “The measures, procedures and remedies provided for in this Directive should be determined in each case in such a manner as to take due account of the specific characteristics of that case, including the specific features of each intellectual property right *and, where appropriate, the intentional or unintentional character of the infringement*” (our italics). It is in this context that proportionality is congruent with the provision of punitive damages.

Thus, we believe the Court is not justified in seeking to view the rule in Polish law as a measure which is merely compensatory: “29. [...] without there being any need to rule on whether or not the introduction of ‘punitive’ damages would

be contrary to Article 13 of Directive 2004/48, it is not evident that the provision applicable in the main proceedings entails an obligation to pay such damages. 30. Thus, it should be pointed out that, where an intellectual property right has been infringed, mere payment of the hypothetical royalty is not capable of guaranteeing compensation in respect of all the loss actually suffered, given that payment of that royalty would not, in itself, ensure reimbursement of any costs – referred to in recital 26 of Directive 2004/48 – that are linked to researching and identifying possible acts of infringement, compensation for possible moral prejudice [...] or payment of interest on the sums due.”

These are *eventual* harms. The fact that proof of the extent of the loss is dispensed with makes it impossible to defend the theory that the amount corresponding to twice the fee is compensatory in nature. Indeed, the costs related to researching and identifying infringements are amounts already included in the amount of the hypothetical fee or, otherwise, they may be compensated according to Article 14 of the Directive as separate damages (“Member States shall ensure that reasonable and proportionate legal costs and other expenses incurred by the successful party shall, as a general rule, be borne by the unsuccessful party, unless equity does not allow this”). Favouring the first statement, see the Conclusions presented by Advocate General Henrik Saugmandsgaard Øe on 4 February 2016 in the case *Jørn Hansson v Jungpflanzen Grünewald GmbH* (Case C-481/14, note 51). Regarding the scope of the costs included in Article 14 the Proposal [COM (2003) 46 final] stated: “Article 18 provides for the legal costs, lawyer’s fees and any other expenses incurred by the successful party (for example, investigation costs, costs for expert opinions) to be borne in full by the other party, unless equity or the economic situation of the other party does not allow this. This possibility is partly provided for by Article 45(2) of the TRIPS Agreement.”

Portuguese law serves as a good example of the use of damages as an instrument to appropriately compensate and, at the same time, repress infringement of intellectual property rights. The costs borne by the injured party to protect the right and to research and bring the infringing conduct to an end are expressly included in the calculation of the compensation (Article 211(2) of the Code of Copyright and Related Rights and Article 338-L(2) of the Industrial Property Code). Besides this, the amount of the costs identified is added to the appropriate royalties or fees, when the use of this criterion proves justified (Article 211(5) of the Code of Copyright and Related Rights and Article 338-L(5) of the Industrial Property Code). When applying the alternative method for compensation, the amount is calculated according to equity and covers, at least, the royalties or fees that would have been paid if the authorisation had been requested and the aforementioned costs.

The laws referred to also allow for all or some of the criteria established for setting the compensation to be combined when the infringer's conduct is a repeated practice or proves to be particularly burdensome (Article 211(6) of the Code of Copyright and Related Rights and Article 338-L(6) of the Industrial Property Code). The Portuguese legislator thus permits, and quite rightly, adjustment of damages to the purposes that Directive 2004/48 seeks, and the proportionality of the penalty must be assessed with reference to the characteristics of the unlawful behaviour and not only the loss⁷.

It will be noted, then, that the rules on compensation cannot make respect for or infringement of the intellectual property rights of others indifferent. The consideration of the profits of the infringer is evidence of this aim. The rules on costs also confirm this (Article 14 – the loser pays principle). It was the dissuasion in question which provided the grounds for associating compensation for moral harm to the lump sum method. On this, Advocate General Melchior Wathelet writes in the Conclusions presented on 19 November 2015 (in the *Liffers* case): “33. [...] it would be inconsistent to exclude compensation for moral prejudice from the damages awarded to a holder of an intellectual property right who chooses to seek compensation for pecuniary loss under the lump-sum method provided for in heading (b) of the second subparagraph of Article 13(1) of Directive 2004/48. 34. Such an exclusion would have the consequence of cancelling out entirely any dissuasive effect of the order made against the infringer, as the latter would be required only to reimburse to the rightholder the sum which he would have been obliged to pay the rightholder if he had respected that right, and that sum may be less than the actual prejudice. Such compensation would not, therefore, be in accordance with the European Union legislature's wish to ensure a high level of protection for intellectual property.” How could a different teleology be justified when the practice is repeated or particularly objectionable? And if the European legislator prefers to omit the solution, it must not prevent the Member States from adopting it.

Highly suggestively, the Advocate General Paolo Mengozzi writes in the Conclusions presented on 3 September 2015, regarding the *Arjona Camacho* case: “47. For the Court to hold unequivocally that Article 18 of Directive 2006/54 is to be interpreted as requiring Member States to provide for the award of punitive damages to victims, in cases of discrimination on grounds of sex which come within the scope of that directive, would be for it to adopt a definitive position in a debate which, to my mind, ought to be left to the Member States. [...] 52. This may well be regrettable, conscious as I am that the system of liability described

⁷ On the subject, see Henrique Sousa Antunes, *Da Inclusão do Lucro Ilícito e de Efeitos Punitivos entre as Consequências da Responsabilidade Civil Extracontratual: a sua Legitimação pelo Dano*, Coimbra, 2011, pp. 327 et seq., particularly pp. 551 et seq.

above is far from fulfilling its compensatory function in a systematically satisfactory way [...].”

We may end by stating that in European law itself the importance of punitive damages is understood, since it is established therein. In this sense, Regulation (EC) No 1768/95 of the Commission, of 24 July 1995, implementing rules on the agricultural exemption provided for in Article 14 (3) of Council Regulation (EC) No 2100/94 on Community plant variety rights, provides in Article 18(2)⁸: “If such person has *repeatedly and intentionally* not complied with his obligation pursuant to Article 14 (3) 4th indent of the basic Regulation, in respect of one or more varieties of the same holder, *the liability to compensate the holder* for any further damage pursuant to Article 94 (2) of the basic Regulation *shall cover at least a lump sum calculated on the basis of the quadruple average amount charged for the licensed production of a corresponding quantity of propagating material of protected varieties of the plant species concerned in the same area, without prejudice to the compensation of any higher damage*” (our italics). The European legislator’s prohibitions on punitive damages in the area of collective protection of rights and competition law are, therefore, *somewhat schizophrenic*.

5. Conclusions

Having undertaken this, necessarily brief, critical assessment of the Judgment, our main conclusions are:

- A. We welcome the Court of Justice’s decision insofar as it recognises that, with regard to compensation, Directive 2004/48 lays down a minimal level of protection;
- B. The Judgment also warrants our agreement when it states that requiring the injured party to prove the damage is contrary to the nature of the compensation provided for in Article 13(1) (b) of the Directive (pre-established damages);
- C. From the analysis of the assumptions on which the decision is based one can see a direction which the Court of Justice, unjustifiably, maintains ambiguous: punitive damages are compatible with Directive 2004/48;
- D. The assessment of proportionality cannot only be based on the extent of the loss, if we accept the legitimacy of punitive damages. In this context, the court should take account of the characteristics of the unlawful act, namely the level of fault of the perpetrator or the repetition of his conduct.

8 In “Official Journal”, L 173, 25.7.1995, pp. 14 to 21.

Die Richtigkeitsgewähr des Tarifvertrags

248 S., Duncker & Humblot, Berlin, 2015

Daniel Krämer

Júlio Gomes

Juiz Conselheiro (STJ)

Professor Convidado da Escola de Direito no Porto
da Universidade Católica Portuguesa

A escolha da presente obra como objeto da recensão resulta não apenas da qualidade da mesma, mas também da circunstância de que uma parte significativa dos temas nela abordados não tem ocupado sobremaneira a doutrina nacional. Ao longo da obra discutem-se, com efeito, não apenas temas já longamente debatidos entre nós como o papel e a relevância da convenção coletiva enquanto fonte de direito, como outros, a saber, de que pressupostos depende o funcionamento da contratação coletiva para que os seus resultados possam ser *prima facie* justos e adequados e qual a sua legitimação¹, qual o significado da autonomia negocial coletiva e se esta está vinculada aos direitos fundamentais de modo análogo ao do legislador, qual a tutela para os direitos individuais, tanto dos filiados, como dos não filiados, e qual a tutela das minorias no processo de negociação coletiva.

Antes de mais, o autor sublinha que historicamente a negociação coletiva surge como reação à insuficiência do contrato individual de trabalho, em face, designadamente, da disparidade de poder entre as partes deste contrato para proporcionar aquela verosimilhança de justiça que SCHMIDT-RIMPLER considerava ser ínsita normalmente no contrato. Pode questionar-se, no entanto, se a disparidade se atenua significativamente só porque há sujeitos coletivos e como o autor refere, de acordo, aliás, com a jurisprudência e a doutrina dominantes na Alemanha, um sindicato só está em uma situação de relativa paridade com um empregador ou uma associação de empregadores se tiver poder social e estiver em condições reais de recorrer a meios de conflito coletivo. Em suma, sindicatos de representatividade praticamente nula não têm, na Alemanha, real capacidade de negociação coletiva, ao contrário do que sucede entre nós em que basta que um sindicato tenha personalidade jurídica para ter capacidade negocial coletiva.

Um outro aspeto destacado pelo autor é o de que a capacidade negocial coletiva não é algo de originário² e referente a uma espécie de direito natural. Associações sindicais e associações de empregadores não devem ser encaradas como entes sociais intermédios dotados de um poder próprio de criação de normas jurídicas, pertencendo a uma espécie de ordenamento jurídico pré-estatal. Bem ao invés, a sua capacidade de criação de normas jurídicas resultaria

1 Como destaca CHRISTIAN ARNOLD, *Betriebliche Tarifnormen und Aussenseiter, Zur Legitimation der tarifvertraglichen Regelungsbefugnis*, Duncker und Humblot, Berlin, 2007, p. 180, refere como problema fundamental da relação entre coletivo e indivíduo, no direito da contratação coletiva, o problema da legitimação da criação de normas através da contratação coletiva que é no fundo também o problema da situação do trabalhador face às associações sindicais, acrescentando que “a legitimação da competência das associações [sindicais] para a criação de normas através da convenção coletiva é um fragmento do problema geral da autorização de uma conduta juridicamente vinculante com eficácia para terceiros” (*ob. cit.*, p. 189).

2 Autor e *ob. cit.*, p. 127.

de uma opção constitucional; é o legislador quem dá às convenções coletivas a possibilidade de serem fontes de direito.

Se até aqui concordamos com o autor, as suas teses, ainda que bem fundamentadas, tornam-se crescentemente controversas a partir destas premissas.

Em primeiro lugar existiria aqui, segundo o autor, uma delegação de poderes normativos do legislador aos sujeitos da negociação coletiva e seria por força da própria vontade do legislador que mesmo os filiados nas associações sindicais ou nas associações de empregadores, estariam abrangidos pelo poder normativo (delegado) da negociação coletiva.

O autor procura realizar aqui, quanto a nós, quase uma quadratura do círculo. Por um lado, as associações sindicais e as associações de empregadores são hoje, no sistema alemão como no nosso, pessoas coletivas de direito privado assentes na livre e espontânea filiação dos seus associados. Mas para o autor a livre filiação é o pressuposto da eficácia normativa (no fim de contas, se os trabalhadores não decidissem filiar-se em sindicatos, decisão que no nosso sistema jurídico é livre, não haveria sindicatos), mas não é o genuíno fundamento da sujeição àquela eficácia normativa de uma convenção. Em suma, o autor nega que trabalhador ou empregador ao filiarem-se expressem uma qualquer sujeição ao poder normativo da associação (ou melhor, ao poder normativo das convenções celebradas por aquela associação com outros sujeitos legitimados a celebrarem convenções coletivas). Se o filiado está sujeito ao poder normativo da associação é apenas por força da vontade do legislador³.

Daí decorre para o Autor que as normas resultantes da convenção coletiva são, mesmo para o filiado em sindicato outorgante, heterónomas⁴. A autonomia coletiva esgota-se na autonomia dos sujeitos ou partes da convenção e a autodeterminação dos filiados resume-se ao ato da filiação e aos seus direitos de eleição e de influência na vida da associação.

Em consequência, e tal como, aliás, boa parte da doutrina, o autor rejeita qualquer apelo à representação civil no contexto da negociação coletiva. Com efeito, afirma, mesmo, que a criação de normas através das convenções coleti-

3 Em termos próximos, na doutrina italiana, mas falando mesmo de um poder originário das associações sindicais cfr., por exemplo, MATTIA PERSIANI, *Osservazioni sulla revisione della dottrina del diritto sindacale*, ADL2011, pp.1 e ss., pp. 5-6: “O poder de exercer a autonomia privada coletiva mediante a celebração do contrato coletivo não pode ser conferido pelos trabalhadores individuais com a sua filiação sindical. E isto porque os trabalhadores individuais só têm os poderes da autonomia privada individual e não os poderes da autonomia privada coletiva. E ninguém pode transmitir mais poderes do que aqueles de que é titular. O poder de exercer a autonomia coletiva e, por conseguinte, de celebrar uma convenção coletiva, não deriva do mandato que com a filiação seria conferido pelos trabalhadores ao sindicato, mas é um poder originário. Um poder conferido ao sindicato diretamente pelo ordenamento jurídico sempre que os trabalhadores associando-se criam um sindicato. O que conta, pois, é que um sindicato exista para que o ordenamento o habilite a celebrar convenções coletivas.”

4 *Ob. cit.*, p. 86.

vas se aproxima muito mais do mecanismo de legitimação democrática em uma relação de autoridade/subordinação com os súbditos do que com uma relação entre cidadãos própria dos contratos de direito privado.

A formação da vontade coletiva expressa pela associação estaria, em princípio, sujeita às regras democráticas e ao princípio da maioria, mas a vontade coletiva não se confunde com a vontade dos filiados, representando, quando muito, no caso ideal, uma expressão da vontade da maioria e da sua eventual disposição a algum compromisso. Em todo o caso, e nas palavras do autor, a explicação da eficácia da convenção coletiva pressupõe uma separação conceptual entre a autonomia coletiva como domínio de autodeterminação e auto-gestão das associações e a sua capacidade de criar normas⁵.

O Autor sublinha a importância da democracia interna das associações e do respeito pelos direitos de tendência, mas acaba desassombadamente por reconhecer que essa estrutura democrática não garante nem tutela de maneira adequada os interesses das minorias. Em primeiro lugar, pelo peso da maioria na formação da vontade coletiva. Depois pelo crescente distanciamento entre bases e cúpulas. A participação dos filiados na formação da vontade acaba por ser frequentemente mediata, já que federações e confederações sindicais acabam por ser associações baseadas no princípio da delegação e coordenadas ou dirigidas por associações de representantes. Mesmo nas associações de base o autor reconhece que muitas minorias ou grupos com interesses especiais tendem a ter uma influência muito reduzida não obtendo uma representação real (indica como exemplos trabalhadores a tempo parcial, trabalhadores precários, etc.). Relativamente às associações de empregadores reconhece que a sua estrutura é diversa – qualificando-a de mais complexa e menos transparente – tendo, em regra, muito menos filiados, mas verificando-se, frequentemente, um predomínio dos interesses dos grandes empregadores que têm amiúde uma influência acentuada em matéria de política da contratação coletiva.

O diagnóstico realizado pelo autor acaba, pois, por ser bastante sombrio. Aliás, interroga-se expressamente sobre quais possam ser os momentos de proteção tanto das minorias, como dos interesses individuais de um filiado e acaba por identificar três momentos⁶ que, no entanto, conferem, no seu entender, uma proteção muito limitada. Um primeiro momento seria o do próprio procedimento de formação da vontade coletiva, muito embora a debilidade dessa tutela resulte tanto do peso fundamental atribuído à maioria como da perda de imediação atrás referida à medida que se constroem associações de associações. Um outro momento estaria na liberdade de o filiado abandonar

5 *Ob. cit.*, p. 126.

6 Autor e *ob. cit.*, pp. 143 e ss.

a associação. Mas como o autor reconhece nem essa liberdade permite ao afiliado que o deixou de ser escapar à aplicação de uma convenção coletiva, permanecendo abrangido por ela, no nosso direito como no direito alemão, até ao seu termo de vigência, para já não falar da possibilidade de ser “perseguido” por um mecanismo de extensão. Acresce que essa liberdade se revela ilusória, mormente em casos em que apenas existe no sector de atividade ou na profissão um sindicato forte, pelo que, mesmo existindo pluralismo sindical do ponto de vista jurídico, não existe uma real alternativa. Um terceiro momento respeita à vinculação deste legislador – os sujeitos da negociação coletiva – aos direitos fundamentais em termos que, para muitos autores, não deverão ser materialmente diferentes ou, pelo menos, menos vinculantes, do que se passa em relação ao legislador democrático “clássico”. No entanto, o próprio autor reconhece que esta vinculação, além de muito controversa, pode não ser suficiente para a tutela dos interesses dos indivíduos e das minorias⁷.

Pese embora a coerência interna da obra e a investigação de boa qualidade em que assenta não podemos deixar de confessar que a sua leitura nos deixou muitas inquietações e, perdoe-se-nos a imagem, um certo sabor amargo.

Em primeiro lugar porque a autonomia negocial coletiva tem nesta visão muito pouco de autonomia. Autónomos são apenas os sujeitos da negociação coletiva, mas os resultados normativos da sua atuação não são a expressão da autonomia privada nem sequer dos seus filiados. A representação civil dá lugar à representatividade, muito próxima da política, tanto que para muitos o sindicato é cada vez mais semelhante a um partido político⁸. Para alguns autores, de resto, a associação sindical, por exemplo, acaba por ser, de algum modo, algo similar a um representante legal⁹ (embora porventura não de uma

7 Autor e *ob. cit.*, p. 149.

8 O papel político do sindicato é enfatizado por alguns autores mesmo a respeito da contratação coletiva, argumentando-se que o instrumental de direito privado é incapaz de o apreender. Neste sentido, vejam-se, por todos, as palavras de GIULIO PROSPERETTI, *L'autonomia collettiva e i diritti sindacali*, UTET, Torino, 2010 (3.ª reimpressão de 2013), p. 6. Para este autor “a representação em sentido técnico traduz-se geralmente numa ficção jurídica já que o carácter político que caracteriza o contrato coletivo não se concilia no plano formal com a relação de mandato”, sublinhando que “a eficácia jurídica dos contratos coletivos prescinde da democracia associativa interna dos sindicatos outorgantes”. O autor acrescenta mesmo que “não é a representação legal de uma determinada direção sindical a determinar a validade do acordo, mas antes a sua eficácia reconhecida de facto pelos empregadores e pelos trabalhadores que aceitam a sua disciplina [...] é o poder de facto que deriva das ações de autotutela do sindicato e designadamente das greves que o sindicato consegue declarar e gerir que legitima o sindicato como contraparte nas negociações para a celebração de contratos coletivos. Para o autor, também, “o sindicato assume um papel político cuja máxima grandeza como expressão de autotutela é a greve geral e o sindicato é também titular de funções para normativas, já que a lei reconhece frequentemente às convenções coletivas funções de integração ou de derrogação da norma legal”.

9 Sirvam de exemplo GEORGES BORENFREUND, *Propos sur la représentativité syndicale*, *Droit Social*, 1988, pp. 476 e ss., e 481, n. 30: “forma atípica de representação legal”, JEAN MAURICE VERDIER, *Sur la relation entre représentation et représentativité syndicales (quelques réflexions, rappels, suggestions)*, *Droit Social*,

categoria, como sucedia no regime corporativo e tendo de suportar, apesar de tudo, a possível coexistência de outras associações sindicais com o mesmo âmbito potencial de representação). A negociação coletiva parece ser acolhida por razões fundamentalmente pragmáticas, entre as quais, possivelmente, a de que facilita o cumprimento pelos destinatários (mormente os trabalhadores) das normas assim criadas, dando-lhes a ilusão de que participaram, através dos seus “representantes” na formação das mesmas. E em toda esta construção não se vislumbra resposta à possível tirania do grupo – que alguns autores (mas não DANIEL KRÄMER) consideram mesmo conceptualmente impossível.

A obra que vimos analisando insere-se na orientação que pensamos ser dominante, sem esquecer que para uma visão mais radical a convenção coletiva já não seria sequer em rigor um contrato.

Mas tem-se assistido, nas últimas décadas, a um ressurgimento de uma outra visão¹⁰, que dir-se-ia historicamente superada, mas que tem regressado sobretudo na doutrina germânica e que tem dado mostras de um insuspeitado vigor. Referimo-nos a uma compreensão da convenção coletiva como expressão de uma autonomia coletiva genuinamente privada a qual retira a sua legitimidade das bases e mormente da filiação voluntária.

Em primeiro lugar, sublinhe-se que se trata de uma visão que se reafirma e que revive em reação ao panorama em muitos aspetos preocupante gerado pela doutrina que nega praticamente relevância à filiação espontânea e à natureza privada das associações sindicais e de empregadores.

Com efeito, num mundo em que é cada vez mais real o perigo de as convenções coletivas se transformarem, na sugestiva expressão de RICHARDI, em grandes árvores sem raízes¹¹, é paradoxal que se atribua a sindicatos com uma

1991, pp.6 e ss. (“a representação sindical não é então apenas uma representação social, mas aproxima-se singularmente de uma representação jurídica, mesmo que não possa ser concebida como uma representação contratual assente sobre um mandato, mesmo face aos filiados dos sindicatos outorgantes, mas antes como uma espécie de representação legal”), PHILIPPE LANGLOIS, *Contrat individuel de travail et convention collective; un nouveau cas de représentation*, *Droit Social*, 1975, pp. 283 e ss., p. 285, VALENTE SIMI, *Formazioni sociali e categoria professionale*, RTDPC, 1965, ano XIX, pp. 426 e ss., pp. 455-456 (“os indivíduos podem filiar-se ou não no sindicato, mas em qualquer uma destas hipóteses permanecem sujeitos à disciplina que este crie, enquanto representante da categoria e que é classificada pela doutrina como um próprio e verdadeiro poder de império”), THILO RAMM, *Die Rechtsnatur des Tarifvertrags*, JZ, 1962, pp. 78 e ss. e 82 (refere-se a algo de semelhante ao poder de representação legal dos incapazes embora depois acrescente que seria como que uma modalidade intermédia entre a representação legal e a representação voluntária).

10 Cfr., por exemplo, REINHARD RICHARDI, *Koalitionsgewalt und individuelle Koalitionsfreiheit*, *Archiv des öffentlichen Rechts* 1968, pp. 243 e ss., p. 254: “a autonomia coletiva, na parte que se reporta à liberdade de negociação coletiva, pertence ao direito privado e é uma modalidade ainda que especial da autonomia privada” (“besondere Erscheinungsform der Privatautonomie”) e EDUARD PICKER, *Tarifautonomie – Betriebsautonomie – Privatautonomie*, NZA 2002, pp. 761 e ss.

11 REINHARD RICHARDI, *Verbandsmitgliedschaft und Tarifgeltung als Grundprinzip der Tarifautonomie*, NZA, 2013, pp. 408 e ss. e 408: “A validade de uma convenção coletiva sem filiação assemelha-se a uma árvore

representatividade muito reduzida – ou até como sucede no nosso sistema legal, simplesmente presumida – o poder de através da negociação com empregadores e associações de empregadores criarem normas destinadas a regular as relações individuais de trabalho, normas essas que podem ser desfavoráveis para os trabalhadores por comparação com as normas legais ou até o poder de qualificarem de modo insindicável como mais favorável uma convenção coletiva em relação à anterior que aquela vem substituir. O perigo de restrição coletivista dos direitos e liberdades dos filiados é também ele um perigo real, de pouco ou nada servindo para o afastar a liberdade sindical negativa cada vez mais esvaziada de conteúdo ou reduzida à mera possibilidade de não pagar quotas sindicais.

Por outro lado se o poder de criação de autênticas normas jurídicas é delegado pelo Estado – ou para outros autores objeto de um reconhecimento pelo Estado – a estas associações sindicais ou de empregadores, poderá mesmo o Estado abdicar de qualquer controlo (ou, pelo menos, da maior parte do controlo) sobre o seu exercício, em nome da liberdade de auto-organização destas associações?

A visão “privatista” da convenção coletiva e do próprio sindicato assenta, assim, em várias premissas.

Assim, e desde logo, no papel (ainda) atribuído ao contrato individual de trabalho. Pode dizer-se que a emergência do Direito do Trabalho e a existência de sistemas de negociação coletiva assentam, em grande medida, na verificação da insuficiência do contrato individual para regular de modo considerado socialmente justo a relação de trabalho face à desigualdade da posição das partes do contrato. Mas esta crítica não pode fazer-nos esquecer que a relação de trabalho continua a ser, em princípio, uma relação contratual, uma relação que se baseia no consentimento livre do trabalhador para trabalhar para outrem e juridicamente subordinado a outrem¹². A liberdade do trabalhador pode ter muitas limitações, mas existe e é relevante porque o trabalho subordinado não se confunde com o trabalho de um servo ou de um escravo. E a aplicação à relação de trabalho de alguns princípios contratuais dos contratos de intercâmbio de prestações pode representar uma das primeiras linhas de defesa dos direitos dos trabalhadores – pense-se na importância da retribuição

sem raízes. A sua copa pode ainda dar sombra, mas falta-lhe o fundamento que lhe pode dar durabilidade.”

12 RICHARDI, *ob. cit.*, p. 408: “Nicht der freie Arbeitsvertrag als Prinzip wurde in Frage gestellt, sondern seine Funktionsfähigkeit für eine freiheitliche Ordnung des Arbeitslebens durch den Abschluss von Tarifverträgen gesichert.” Em suma, a autonomia negocial coletiva não é a negação da autonomia individual, mas o seu reforço e a sua garantia.

contratualmente acordada, na proibição em princípio de modificação unilateral do objeto do contrato (ainda que com muitas exceções no direito do trabalho), na proibição de vínculos tendencialmente perpétuos. Mas, e sobretudo, deve conceber-se a autonomia coletiva como um meio de reforçar e não de suprimir a autonomia individual. Trata-se de uma autotutela exercida coletivamente¹³ através de uma organização que é expressão da liberdade dos seus filiados e que deve servir o interesse destes e respeitar um núcleo irredutível de direitos individuais de cada filiado. E o sindicato representa em primeira linha aqueles que nele se filiam e que confiam, através da sua filiação livre, naquele sindicato para a defesa e representação dos seus interesses (individuais e como coletivo). A legitimação destas associações resulta, assim, em primeira linha, segundo esta conceção, da filiação, tratando-se uma legitimação “vertical ascendente”. É certo que – como aliás sucede com um qualquer contrato de direito civil – a comunidade politicamente organizada colocou ao serviço dos particulares meios de coerção – se, em regra, é necessário que a lei defina, por exemplo, o que conta como um contrato e quais as consequências da sua violação, aqui foi necessário definir, por exemplo, quem são os sujeitos que podem celebrar convenções coletivas e quais os efeitos dessas convenções, mormente normativos. Há aqui, pelo menos nos países de *civil law*, um meio de atuação diverso e reforçado relativamente aos outros contratos, porventura mesmo em relação aos chamados “contratos normativos”, de direito privado, através do reconhecimento de uma fonte de direito, meio de atuação que só é possível através do poder do Estado que assumiu a tarefa da promoção da contratação coletiva. Esse meio significa que o sindicato quando celebra uma convenção coletiva não se limita a assumir em nome dos filiados direitos e obrigações contratuais, não se tratando da representação civil clássica: o sindicato é parte, é o sujeito da convenção (e não os seus filiados) e pode negociar muitas matérias que não poderiam sequer ser negociadas ao nível dos contratos individuais de trabalho. Mas seria errado, quanto a nós, conceber os filiados como “terceiros” clássicos totalmente estranhos à convenção e a quem ela se aplica heteronomamente como pretende DANIEL KRÄMER. A lei – desde logo a Constituição e depois a lei ordinária – ao criar o mecanismo da negociação coletiva abre um novo espaço para o exercício pelos trabalhadores da sua autonomia privada, embora um espaço necessariamente coletivo. Associando-se livre e espontaneamente em sindicatos podem participar na negociação de matérias que estariam vedadas

13 Quanto a este aspeto é ainda hoje fundamental a obra de ROSARIO FLAMMIA, *Contributo all'Analisi dei Sindacati di Fatto*, I – Autotutela degli Interessi di Lavoro, Giuffrè, Milano, 1963. Muito embora o autor considerasse que a autotutela coletiva e sindical não se reconduzia à autonomia privada, não deixava de afirmar que “as formas privatistas são, no meu entender, as únicas que correspondem à realidade sindical própria de um ordenamento jurídico democrático do tipo neocapitalista”.

à contratação individual e na criação de normas que irão reger mesmo aspetos sobre os quais pode incidir o contrato individual de trabalho. Mas o espaço criado não deixa de ser um espaço de autonomia e de liberdade negociais¹⁴.

Esta visão tem, evidentemente, uma série de corolários. Por razões de espaço, iremos referir apenas alguns. Nesta visão surge como normal – mas também extremamente relevante o princípio do tratamento mais favorável e a prevalência do contrato individual de trabalho mais favorável sobre a convenção coletiva¹⁵. Por outro lado, a consciência de que através da negociação coletiva as partes procuram satisfazer interesses coletivos¹⁶, mas estes, apesar da sua relevância, não se confundem com o bem comum¹⁷ que apenas o soberano democraticamente eleito prossegue. Depois, que importa considerar que estas pessoas privadas, que têm o poder de criar normas jurídicas, devem estar vinculadas ao respeito pelos direitos fundamentais de modo análogo ou muito similar ao modo de vinculação do legislador. E, ainda, que existe uma margem irreduzível de liberdade individual, em suma, matérias em que a dimensão individual é tão importante que deve ter-se por excluída uma intervenção da convenção coletiva sobretudo quando esta reduz ou cerceia direitos conferidos por lei.

14 Contra visualizando a autonomia coletiva como sendo simultaneamente uma limitação da autonomia privada dos empregadores e dos trabalhadores na regulamentação da sua vida. Cfr., por exemplo, DIETER REUTER, *Möglichkeiten und Grenzen einer Deregulierung des Arbeitsrechts*, Festschrift für Herbert Wiedemann zum 70. Geburtstag, Verlag C. H. Beck, München, 2002, pp. 449 e 453.

15 Sobre o tema cfr., por todos, MEINHARD HEINZE, *Tarifautonomie und sogennantes Günstigkeitsprinzip*, NZA, 1991, pp. 329 e ss. O autor sublinha que o princípio do tratamento mais favorável, mais concretamente a prevalência do contrato de trabalho individual mais favorável sobre o contrato coletivo não deve ser discutida isoladamente (p. 330). Na verdade deve discutir-se se e em que medida é que a autonomia privada e os direitos individuais dos trabalhadores não colocam também limites à autonomia coletiva. A prevalência do contrato individual de trabalho mais favorável demonstra que a escolha autónoma do trabalhador individual não é apenas consentida, mas é também válida. É na própria liberdade individual que se fundamenta também o associativismo livre. O autor considera que seria inconstitucional porque violadora da liberdade da pessoa, uma visão das associações sindicais que negasse a liberdade contratual do indivíduo, absolutizando a liberdade contratual do coletivo. A autonomia negocial coletiva tem uma função instrumental, visando reforçar a própria liberdade individual. Além disso, a autonomia negocial coletiva não pode monopolizar a configuração das relações de trabalho sob pena de o direito ao contrato individual e à autonomia privada individual serem completamente esvaziados. Alguns autores defendem, por isso, que a intervenção da autonomia coletiva deveria por isso ser de algum modo subsidiária – cfr., por exemplo, IVO NATZEL, *Subsidiaritätsprinzip im kollektiven Arbeitsrecht*, Zeitschrift für Arbeitsrecht, 2003, pp. 103 e ss.

16 REINHARD RICHARDI, *Koalitionsgewalt und individuelle Koalitionsfreiheit*, Archiv des öffentlichen Rechts, 1968, pp. 243 e ss. e 254: As partes da convenção coletiva não realizam tarefas do Estado, mas zelam e defendem os interesses dos seus filiados.

17 Como acentua VOLKER RIEBLE, *Der Tarifvertrag als kollektiv-privatautonomer Vertrag*, ZfA, 2000, pp. 5 e ss., o interesse coletivo prosseguido por estas associações não se confunde com o bem comum. O autor denuncia o que considera ser neste contexto um “misticismo coletivista”.

DIREITO
PRIVADO

Carlos Ferreira de Almeida
Joseph William Singer
Pablo Mañá Pinto
Henrique Soares Arturios
Júlio Gomes



Revista Quadrimestral

- Sim, desejo ser assinante a partir do próximo número.
Ficha para assinatura em www.fd.lisboa.ucp.pt

- Portugal** Assinatura anual 36,00 €
Europa Assinatura anual 46,00 €
Resto do Mundo Assinatura anual 56,00 €

Nome
Morada
Localidade Código Postal
NIF* Tel. Tm.
e-mail Profissão
Instituição | Empresa

* Campo de preenchimento obrigatório para emissão de recibo.

FORMAS DE PAGAMENTO

- Cheque n.º
Banco
Cheque passado a:
Universidade Católica Portuguesa
 Transferência bancária
PT50 0033 0000 2108 0018 4224 3
(Anexar comprovativo de transferência)

ASSINATURA

- OS DADOS FORNECIDOS SERÃO PROCESSADOS AUTOMATICAMENTE NO FICHEIRO COMERCIAL DO DESTINATÁRIO, GOZANDO OS RESPECTIVOS TITULARES DE DIREITO DE INFORMAÇÃO E ACESSO. ESTAS INFORMAÇÕES SERÃO EXCLUSIVAMENTE UTILIZADAS PARA DIVULGAÇÃO DE PRODUTOS E PROMOÇÕES DA CATÓLICA LAW REVIEW. SE, DE FUTURO, NÃO PRETENDE RECEBER ESSAS INFORMAÇÕES, ASSINALE UM X.

POR FAVOR, ENVIE ESTE CUPÃO PARA:

Católica Research Centre for the Future of Law
Faculdade de Direito
Universidade Católica Portuguesa
Palma de Cima 1649-023 Lisboa

INFORMAÇÕES: Tel.: (351) 21 721 41 78 | Fax (351) 21 721 41 77
catolicalawreview.fd@ucp.fd.pt



VOLUME I
n.º 2
mai. 2017

DIREITO PRIVADO

DOCTRINA

Invalidez, inexistência e ineficácia \ **Carlos Ferreira de Almeida**

We don't serve your kind here:
public Accommodations and the mark of Sodom \ **Joseph William Singer**

Sobre a alegada "Superação" do Direito pela análise económica \ **Paulo Mota Pinto**

COMENTÁRIO DE JURISPRUDÊNCIA

Intellectual property rights:
infringement and calculation of damages \ **Henrique Sousa Antunes**

RECENSÃO

Die Richtigkeitsgewähr des Tarifvertrags \ **Júlio Gomes**



FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

Disponível em e-book