通过编制中国2019-2023年间153个部门的投入产出表，并使用了分数多项Logit（FMNL）模型，分析能源部门结构变化对投入产出系数的影响。

1. **数据预处理**

首先从国家统计局、产业数据库等多来源收集各部门总产值、增加值、最终使用和进口的相关数据，对数据进行整合估计，以作为后续分析的输入。估计方法主要参考《中国时间序列投入产出表的编制: 1981—2018》（张红霞 et al., 2021）。

1. **总产值的确定**

由于除农林牧渔业外的其他产业部门并没有现成的总产值数据，或统计口径与投入产出表要求的并不一致（如统计年鉴中给出的建筑业总产值其实是资质内建筑业总产值），编制中采用的方法为：

其中参考指标主要选用与总产值联系密切的统计指标，如工业部门采用规模以上工业企业营业收入，批发零售部门采用限额内营业收入（营利性服务业以经营活动的成果衡量），运输业采用客运/货运量，卫生、教育等行业采用从业人员工资总数（非营利性服务业以费用衡量）等。

1. **增加值的确定**

统计年鉴提供了部分部门增加值数据，但细分程度不够。编制过程中主要采用总量+结构的估计方法，即采用合适的价格指数对参考年度的各部门增加值进行调整，作为编表年度各部门增加值的结构参考，再使用已知的总量数据为各部门分配增加值。

对第一、二、三产业，分别参考供给表数据对已知个大类的增加值数据进行初步的细分，再进行结构调整，分配给该产业对应投入产出部门的个部门。如对于第二产业，统计年鉴提供了采矿业、制造业、电热燃气水生产供应业三个大类的增加值数据，则，而投入产出表中第二产业的部门有97个，则。

第个产业的各部门增加值确定方法如下：

1. 从统计年鉴得到分大类的增加值数据
2. 从2018年的供给表中得到产出系数矩阵 ，即同一产业生产的不同产品占该产业总产出的比例
3. 得到细分产业的增加值数据

对于工业部门，在上述方法基础上，利用统计局披露的月度数据中工业部门不变价增加值的增长速度及经济普查年份的较为细分的工业部门增加值数据，进一步调整各部门的增加值。

1. **最终使用和进口的确定**

对于最终使用列，由于分部门最终使用的数据基础不佳，采用支出法生产总值的增长调整参考年度的各部门最终使用，作为最终使用列的初始值。

对于进口列，基于编表年份相对参考年份分类别进口商品总额的增长，调整参考年度的各部门进口，作为编表年份进口列的初始值。

接着，使用R语言的Nloptr包，利用COBYLA算法最小化确定值与初始值的绝对离差之和，来优化进口和最终使用数据，以确保特定行业和整体经济的最终使用和进口数据满足一致性约束，并尽可能减少调整幅度。 COBYLA算法是一种用于解决有约束非线性优化问题的优化算法，它适用于解决那些目标函数或约束条件较为复杂、难以通过传统的解析方法求解的问题。COBYLA算法通过线性逼近的方法来处理约束条件，从而能够在一定程度上避免对目标函数或约束条件的复杂求导过程。

算法主要实现方法如下：

1. 针对特定行业进行处理，由参考年份的投入产出表知某些行业的最终使用和进口数据满足中间使用合计为0的约束条件，即。‌**‌**‌如住宅房屋建筑、社会工作等部门。
2. 对剩余行业进行全局调整，优化过程中不仅需要所有部门调整后的最终使用和进口的合计满足等式约束，即，还要求它们满足额外的不等式约束，即调整后的数据在给定的上下限内变化，并且根据调整后数值计算的各部门中间使用应与参考年的使用量的比值需要保持在合理的区间内。
3. 整个优化过程通过逐步调整最终使用和进口数据，并在每次迭代中监控目标函数和约束的变化，确保最终结果既满足各行业的一致性要求，又最大限度地减少调整对原始数据的偏离。
4. **估计中间流量矩阵**

通过重点系数法与GRAS（Generalized RAS）相结合的方法，估计中间流量矩阵。估计方法主要参考国家统计局公布的我国投入产出延长表编制方法。

**a. 生成初始直接消耗矩阵**

从参考年的投入产出表中提取数据并计算参考年的直接消耗矩阵，对各部门进行相应的价格调整，得到编表年的初始直接消耗矩阵。直接消耗矩阵表示了每个部门的投入与总产出的比例关系。

**b. 确定重点系数矩阵并计算重点流量矩阵**

使用70%的规则来识别重点系数。即在行方向上，某部门中间使用合计占比70%及以上的部分，以及在列方向上，某部门中间投入合计占比70%及以上的部分，都被标记为重点系数。对于重点系数，将对应该部门的值赋予重点系数矩阵对应的位置，其他元素为0。

利用重点系数矩阵右乘各部门总产出，得到重点流量矩阵。

**c. 生成初始非重点系数矩阵和初始非重点流量矩阵**

初始非重点系数矩阵是在直接消耗矩阵基础上，将重点系数位置的值取为0，而对非重点系数的值进行保留得到的。

计算初始非重点流量矩阵，该矩阵同样是通过乘以总产出矩阵得到，即。

**d. 使用GRAS方法迭代更新得到非重点流量矩阵**

首先定义GRAS算法。算法的定义参考文献《The Solution of Updating or Regionalizing a Matrix with both Positive and Negative Entries》(Junius & Oosterhaven, 2003)。GRAS算法中包含了矩阵的初始定义、迭代计算和精度检验，目标函数是使得调整矩阵相对初始矩阵的信息损耗最小。具体逻辑如下：

1. 定义输入矩阵、目标行、列控制向量、最大迭代次数、精度要求等参数。根据这些输入，将原始矩阵分解为非负部分和非正部分。同时初始化两个对角矩阵，分别用于行和列的缩放。这些对角矩阵将在后续的迭代过程中逐步调整，以逼近目标行和列总和。
2. 通过循环来逐步调整行和列的缩放因子。

列的更新：计算当前列对应的非负部分和非正部分的加权和。根据这些加权和以及目标列总和，更新列的缩放因子。为了避免数值不稳定，需要对缩放因子进行了边界处理，如果因子接近于零或出现了无效值，则设置一个最小值。

行的更新：与列更新类似，计算当前行的非负部分和非正部分的加权和。然后，依据目标行总和更新行的缩放因子，同样地，行缩放因子也被设置了最小值，以防止数值问题。

1. 在每次迭代结束时，计算当前矩阵与目标矩阵之间的误差，并与给定的精度要求进行比较。如果误差小于预设的精度，则停止迭代，否则继续进行下一轮的行列更新。
2. 当迭代结束后，生成一个调整后的矩阵。这个矩阵的生成基于调整后的行和列缩放因子，非负部分的元素被放大，而非正部分的元素则被缩小，确保了矩阵元素被调整到符合新的行列总和的同时，保持与原始矩阵结构的相似性。为了确保调整后的矩阵不会偏离原始值过多，对每个元素的调整值进行限制，使其保持在原始值的±20%范围内。
3. 再次检查调整后的矩阵是否满足目标行列总和。如果某一行或某一列的总和与目标值不符，函数将重新对该行或列进行缩放调整，直到满足要求为止。最终，函数返回满足所有约束条件的调整后的矩阵。

接着根据输入的数据，计算行控制量和列控制量，用来约束非重点系数矩阵的迭代调整。

最后使用GRAS算法调整非重点流量矩阵，使其符合行列控制量的约束条件。GRAS算法迭代更新矩阵和，直到达到指定的精度为止。

**e. 合成中间流量矩阵**

最终的中间流量矩阵是通过重点流量矩阵与调整后的非重点流量矩阵相加得到的，即。

1. **构建FMNL模型**

模型的构建方法主要参考文献《Predicting Structural Changes of the Energy Sector in an Input–output Framework》(Papke & Wooldridge, 1996; Wimmer et al., 2023)。

1. **读取数据**

定义read\_province\_data函数，读取指定年份的所有省级投入产出表文件，并根据不同年份和省份应用不同的读取范围，在读取过程中同步计算各个投入产出表对应的投入产出系数。其中，对于2017年及之后的年份，定义turn\_into\_42sectors函数，将读取到的数据从原始的多个部门合并为42个部门。2019、2021、2022、2023年的投入产出表由第一、二部分的方法估计生成。读取的数据根据年份和省份索引被存储在一个列表中，以便后续使用。

接着，从已生成的储存各省份各年的能源消耗占比数据文件中，循环遍历每一行数据，并与第一步生成的投入产出系数及相关的控制变量合并，来生成输入模型的数据集。

1. **拟合模型**

分数多项式（Fractional Multinominal Logit, FMNL）模型是多项式 Logit 模型的扩展。标准的多项式 Logit 模型只考虑 0-1 的响应变量，而分数多项式 Logit 模型则考虑了响应变量为分数且总和为 1 的情况。

STATA和R均提供fmlogit包，沿用了Papke and Wooldridge (1996) 提出的方法，使用拟极大似然估计（quasi-maximum likelihood）估计FMNL模型的方差渐近一致的参数，并提供了稳健标准误差。

使用STATA的fmlogit包，分别将投入系数、使用系数作为因变量构建FMNL模型。其中，每一个模型均使用了人均GDP、能源消耗总量、人口数作为控制变量，并控制了个体和时间的固定效应。

使用2020年和2023年的输入变量进行拟合效果的检验，发现拟合得到的87个系数与真实值的差值的绝对值平均为 。其中，没有发现任何一个部门的实际系数与估计系数之间存在巨大差异，这表明我们的模型合理地拟合了当前的响应变量。

1. **估计新结构下的中间流量矩阵**

在对能源使用成本进行拆分后，可以得到10%的电力总投入从煤电转移至5%的风电和5%的光电的情景下，电力、热力的生产和供应部门的中间投入列向量的变化情况。

在第三部分得到的模型基础上，预测2023年能源结构调整后的投入产出系数。控制其他变量不变，假设2023年同样有10%的煤电转移至风光电。将新的各种能源占比输入模型，即可预测2023年能源结构调整后的投入产出系数。

使用与第二部分类似的方法，在确定了2023年重点系数后，额外控制了电力、热力的生产和供应部门的系数，将其余的元素作为非重点系数输入到GRAS算法中进行迭代调整，最终可得对应新的投入产出系数的中间流量矩阵。